

Promotoren zur Expression von Genen in Tagetes

Beschreibung

- 5 Die vorliegende Erfindung betrifft die Verwendung von Promotoren zur Expression, vorzugsweise zur blütenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes, die genetisch veränderten Pflanzen der Gattung Tagetes sowie ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung der genetisch veränderten Pflanzen.
- 10 Verschiedene biosynthetische Produkte, wie beispielsweise Feinchemikalien, wie unter anderem Aminosäuren, Vitamine, Carotinoide, aber auch Proteine werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik-, Feed-, Food- und pharmazeutischen Industrie.
- 15 Diese Substanzen, die zusammen als Feinchemikalien/Proteine bezeichnet werden, umfassen unter anderem organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Nukleotide und Nukleoside, Lipide und Fettsäuren, Diole,
- 20 Kohlenhydrate, aromatische Verbindungen, Vitamine, Carotinoide und Cofaktoren, sowie Proteine und Enzyme. Ihre Produktion im Großmaßstab erfolgt zum Teil mittels biotechnologischer Verfahren unter Verwendung von Mikroorganismen, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren.
- 25 Carotinoide werden de novo in Bakterien, Algen, Pilzen und Pflanzen synthetisiert. In den letzten Jahren wird zunehmend versucht, auch Pflanzen als Produktionsorganismen für Feinchemikalien, insbesondere für Vitamine und Carotinoide zu nutzen.
- 30 Ein natürliches Gemisch aus den Carotinoiden Lutein und Zeaxanthin wird beispielsweise aus den Blüten von Marigold Pflanzen (Tagetes Pflanzen) als sogenanntes Oleoresin extrahiert. Diese Oleoresin findet Anwendung sowohl als Inhaltsstoff von Nahrungsergänzungsmitteln als auch im Feed-Bereich.
- 35 Lycopin aus Tomaten findet ebenso Anwendung als Nahrungsergänzungsmittel, während Phytoen überwiegend im kosmetischen Bereich verwendet wird.
- Ketocarotinoide, also Carotinoide, die mindestens eine Keto-Gruppe enthalten, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-
- 40 Hydroxyechinenon, Adonirubin und Adonixanthin sind natürliche Antioxidantien und

Pigmente, die von einigen Algen, Pflanzen und Mikroorganismen als Sekundärmetabolite produziert werden.

Aufgrund ihrer farbgebenden Eigenschaften werden die Ketocarotinoide und insbesondere Astaxanthin als Pigmentierhilfsstoffe in der Tierernährung, insbesondere in der Forellen-, Lachs- und Shrimpszucht verwendet.

Ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von natürlichen, biosynthetischen Produkten und insbesondere Carotinoiden ist daher von großer Bedeutung.

WO 0032788 beschreibt einige Carotinoid Biosynthesegene aus Pflanzen der Gattung Tagetes und offenbart, wie genetisch veränderte Pflanzen der Gattung Tagetes hergestellt werden könnten, um in den Petalen verschiedene Carotinoidprofile zu erhalten und damit gezielt bestimmte Carotinoide herzustellen. Dazu ist es nötig, einige Biosynthesegene überzuexprimieren und andere zu unterdrücken.

Zur Überexpression der neu gefundenen Carotinoid-Biosynthesegene in Pflanzen der Gattung Tagetes wird in WO 0032788 der petalenspezifische Promotor der Ketolase aus Adonis vernalis postuliert.

Aufgrund einer Vielzahl möglicher Schwierigkeiten bei der Überexpression bestimmter Gene besteht ein ständiges Bedürfnis, weitere Promotoren zur Verfügung zu stellen, die eine Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.

Der Erfindung lag daher die Aufgabe zu Grunde, weitere Promotoren zur Verfügung zu stellen, die die Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.

Demgemäß wurde gefunden, dass sich die Promotoren, ausgewählt aus der Gruppe

- 30 A) EPSPS Promotor
- B) B-Gene Promotor
- C) PDS Promotor und
- D) CHRC Promotor

35 sehr gut zur Expression von Genen in Pflanzen der Gattung Tagetes eignen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

Die Erfindung betrifft daher die Verwendung eines Promotors, ausgewählt aus der Gruppe

- 5 A) EPSPS Promotor
B) B-Gene Promotor
C) PDS Promotor und
D) CHRC Promotor

zur Expression von Genen in Pflanzen der Gattung *Tagetes*, mit der Maßgabe, dass Gene aus Pflanzen der Gattung *Tagetes*, die in Wildtyppflanzen der Gattung *Tagetes* von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

15 Benfey et al. (Plant Cell Volume 2, pp. 849-856) beschreiben den EPSPS Promotor aus Petunia als petalenspezifischen Promotor zur Expression von Genen in Petunia hybrida.

Ronen et al. (PNAS Volume 97, Number 20, 11102-11107 beschreiben den B-GENE Promotor aus Tomate als blütenspezifischen Promotor zur Expression von Genen in Tomaten.

Corona et al. (Plant Journal Volume 9 Number 4 pp. 505-512), Mann et al. (Nature Biotechnology Volume 18 pp. 888-892) und Rosati et al. (Plant Journal Volume 24 Number 3 413-419) beschreiben den PDS Promotor aus Tomate als frucht- und blütenspezifischen Promotor zur Expression von Genen in Tomaten und Tabak.

25 Vishnevetsky et al. (Plant Journal Volume 20 Number 4 pp. 423-431) beschreiben den CHRC Promotor aus Gurke als blütenspezifischen Promotor zur Expression von Genen in Gurke, und weiteren Pflanzen wie z.B. Nelke, Sonnenblume, Tabak.

30 Es sind weiterhin zahlreiche blütenspezifische Promotoren aus verschiedenen Organismen in der Literatur bekannt. Dabei wurde überraschend festgestellt, dass viele dieser Promotoren in Pflanzen der Gattung Tagetes nicht zur Expression, insbesondere nicht zur blütenspezifischen oder petalenspezifischen Expression von Genen führen.

35 Es war daher überraschend, dass sich die Promotoren, ausgewählt aus der Gruppe

- A) EPSPS Promotor
 - B) B-Gene Promotor

- C) PDS Promotor und
- D) CHRC Promotor

5 sehr gut zur Expression, insbesondere zur blütenspezifischen und besonderes bevorzugt zur petalenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes eignen.

10 Unter einem Promotor wird erfindungsgemäß eine Nukleinsäure mit Expressionsaktivität verstanden, also eine Nukleinsäure verstanden, die in funktioneller Verknüpfung mit einer zu exprimierenden Nukleinsäure, im folgenden auch Gen bezeichnet, die Expression, also die Transkription und die Translation dieser Nukleinsäure oder dieses Gens reguliert.

15 Unter „Transkription“ wird erfindungsgemäß der Prozess verstanden, durch den ausgehend von einer DNA-Matrize ein komplementäres RNA-Molekül hergestellt wird. An diesem Prozess sind Proteine wie die RNA-Polymerase, sogenannte Sigma-Faktoren und transkriptionelle Regulatorproteine beteiligt. Die synthetisierte RNA dient dann als Matrize im Prozess der Translation, der dann zum biosynthetisch aktiven Protein führt.

20 Unter einer „funktionellen Verknüpfung“ versteht man in diesem Zusammenhang beispielsweise die sequentielle Anordnung einer der erfindungsgemäßen Promotoren und einer zu exprimierenden Nukleinsäuresequenz und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der Nukleinsäuresequenz erfüllen kann. Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu exprimierende Nukleinsäuresequenz oder das zu exprimierende Gen hinter (d.h. am 3'-Ende) der erfindungsgemäßen Promotorsequenz positioniert wird, so dass beide Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

35 Unter „Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate, verstanden.

Unter „spezifischer Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

- Bei einer „verursachten Expressionsaktivität“ oder „verursachten Expressionsrate“ im
5 Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp so nicht vorhanden war.

- Bei einer „erhöhten Expressionsaktivität“ oder „erhöhten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp in einer
10 bestimmten Zeit die gebildete Menge des Proteins erhöht.

Die Bildungsrate, mit der ein biosynthetisch aktives Protein hergestellt wird, ist ein Produkt aus der Rate der Transkription und der Translation. Beide Raten können erfindungsgemäß beeinflusst werden und damit die Rate der Bildung von Produkten in ei-
15 nem Mikroorganismus beeinflussen.

- Die Bezeichnung „dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem „jeweiligen“ Promotor exprimiert werden, ausgenommen sind“, bedeutet, dass beispielsweise der EPSPS Promotor aus Pflanzen der
20 Gattung Tagetes nicht zur Expression von EPSPS-Genen aus Pflanzen der Gattung Tagetes verwendet wird. Dahingegen kann das EPSPS-Gen aus Pflanzen der Gattung Tagetes erfindungsgemäß durch einen B-Gene Promotor, PDS Promotor oder CHRC Promotor aus Pflanzen der Gattung Tagetes exprimiert werden.

- 25 Unter dem Begriff "Wildtyp" oder „Wildtyppflanze“ wird erfindungsgemäß die entsprechende Ausgangspflanze der Gattung Tagetes verstanden.

- Je nach Zusammenhang kann unter dem Begriff "Pflanze" die Ausgangspflanze (Wildtyp) oder eine erfindungsgemäße, genetisch veränderte Pflanze der Gattung Tagetes
30 oder beides verstanden werden.

- Vorzugsweise wird unter "Wildtyp" für die Erhöhung oder Verursachung der Expressionsaktivität oder Expressionsrate und für die Erhöhung des Gehalts an biosynthetischen Produkten die Pflanze *Tagetes erecta*, insbesondere die Pflanze *Tagetes erecta*
35 Hybrid 50011 (WO 02012438) als Referenzorganismus verstanden.

- Unter einem „EPSPS Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine 5-Enolpyruylshikimat-3-phosphatsynthase, regulieren, sowie von diesen
40 Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder

durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

- 5 Diese EPSPS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen EPSPS Promotorsequenzen oder den Nukleinsäuren, kodierend eine 5-
10 Enolpyruylshikimat-3-phosphatsynthase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine 5-Enolpyruylshikimat-3-phosphatsynthase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

- 15 Unter einer 5-Enolpyruylshikimat-3-phosphatsynthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Shikimat-3-Phosphat in 5-Enolpyruylshikimat-3-Phosphat umzuwandeln.

- 20 Bevorzugte EPSPS Promotoren enthalten

- A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder
A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von
Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 %
25 auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder
3 aufweist oder
A3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.
1, 2 oder 3 unter stringenten Bedingungen hybridisiert oder
A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3)

- 30
Die Nukleinsäuresequenz SEQ. ID. NO. 1 stellt eine Promotorsequenz der 5-
Enolpyruylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida (AAH19653)
dar.
35 Die Nukleinsäuresequenz SEQ. ID. NO. 2 stellt eine Promotorsequenz der 5-
Enolpyruylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida (M37029) dar.
40 Die Nukleinsäuresequenz SEQ. ID. NO. 3 stellt eine weitere Promotorsequenz der 5-
Enolpyruylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida dar.

Die Erfindung betrifft weiterhin EPSPS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 1, 2 oder 3) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nuklein-
5 säureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1,2 oder 3 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße EPSPS Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren 10 genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 1, 2 oder 3 leicht auffinden.

Künstliche erfindungsgemäße EPSPS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3 durch künstliche Variation und Mutation, 15 beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auf- finden.

Die folgenden Definition und Bedingungen der Identitätsvergleiche und Hybridisie-
20 rungsbedingungen gelten für alle Nukleinsäuren, also alle Promotoren und Gene der Beschreibung.

Unter dem Begriff "Substitution" ist der Austausch einer oder mehrerer Nukleotide durch ein oder mehrere Nukleotide zu verstehen. „Deletion“ ist das Ersetzen eines Nukleotides durch eine direkte Bindung. Insertionen sind Einfügungen von Nukleotiden 25 in die Nukleinsäuresequenz, wobei formal eine direkte Bindung durch ein oder mehrere Nukleotide ersetzt wird.

Unter Identität zwischen zwei Nukleinsäuren wird die Identität der Nukleotide über die jeweils gesamte Nukleinsäurelänge verstanden, insbesondere die Identität die durch 30 Vergleich mit Hilfe der Vector NTI Suite 7.1 Software der Firma Informax (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

35 Multiple alignment parameter:
Gap opening penalty 10
Gap extension penalty 10
Gap separation penalty range 8
Gap separation penalty off
40 % identity for alignment delay 40

Residue specific gaps off
Hydrophilic residue gap off
Transition weighing 0

5 Pairwise alignment parameter:

FAST algorithm on

K-tuplesize 1

Gap penalty 3

Window size 5

10 Number of best diagonals 5

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 1 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 1, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 2 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 2, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 3 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 3, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

30 Besonders bevorzugte EPSPS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

35 Weitere natürliche Beispiele für EPSPS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3, aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

40

Ein weiterer Gegenstand der Erfindung betrifft daher EPSPS Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150

5 Nukleotide.

Unter "hybridisieren" versteht man die Fähigkeit eines Poly- oder Oligonukleotids, unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären

10 Partnern unterbleiben. Dazu sollten die Sequenzen vorzugsweise zu 90-100% komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze.

15 Die Hybridisierung erfolgt erfinungsgemäß unter stringenten Bedingungen. Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben:

20 Unter stringenten Hybridisierungs-Bedingungen werden insbesondere verstanden: Die über Nacht Inkubation bei 42°C in einer Lösung bestehend aus 50 % Formamid, 5 x SSC (750 mM NaCl, 75 mM Tri-Natrium Citrat), 50 mM Natrium Phosphat (ph7,6), 5x Denhardt Lösung, 10% Dextranulfat und 20 g/ml denaturierte, gescheerte
25 Lachsspermien-DNA, gefolgt von einem Waschen der Filter mit 0,1x SSC bei 65°C.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

30 Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

35 Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform A1), A2) oder A3) beschriebenen EPSPS Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ.

ID. NO. 1, 2 oder 3 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 als EPSPS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung

5 Tagetes.

Alle vorstehend erwähnten EPSPS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäure-

10 bausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren

15 werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem „B-Gene Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Lycopin- β -Cyclase, insbesondere eine chromoplastenspezifische Lycopin- β -Cyclase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

25 Diese B-Gene Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen B-Gene Promotorsequenzen oder den Nukleinsäuren, kodierend eine Lycopin- β -Cyclase, auffinden.

30 Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Lycopin- β -Cyclase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

35 Unter einer Lycopin- β -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Lycopin in γ -Carotin und/oder β -Carotin umzuwandeln.

Bevorzugte B-Gene Promotoren enthalten

- B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder
B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von
5 Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 %
auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder
6 aufweist oder
B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.
10 4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder
B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3)

Die Nukleinsäuresequenz SEQ. ID. NO. 4 stellt eine Promotorsequenz der chro-
moplastenspezifischen Lycopin-β-Cyclase (B-Gene) aus *Lycopersicon esculentum*
(AAZ51517) dar.

15 Die Nukleinsäuresequenz SEQ. ID. NO. 5 stellt eine Promotorsequenz der chro-
moplastenspezifischen Lycopin-β-Cyclase (B-Gene) aus *Lycopersicon esculentum*
(AAZ51521) dar.

20 Die Nukleinsäuresequenz SEQ. ID. NO. 6 stellt eine weitere Promotorsequenz der
chromoplastenspezifischen Lycopin-β-Cyclase (B-Gene) aus *Lycopersicon esculentum*
dar.

25 Die Erfindung betrifft weiterhin B-Gene Promotoren, enthaltend eine von diesen Se-
quenzen (SEQ. ID. NO. 4, 5 oder 6) durch Substitution, Insertion oder Deletion von
Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nuklein-
säureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist.

30 Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße B-Gene
Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren
genomische Sequenz bekannt ist, durch Identitätsvergleiche der
Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen
Sequenzen SEQ ID NO: 4, 5 oder 6 leicht auffinden.

35 Künstliche erfindungsgemäße B-Gene Promotor-Sequenzen lassen sich ausgehend
von den Sequenzen SEQ ID NO: 4, 5 oder 6 durch künstliche Variation und Mutation,
beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auf-
finden.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 4 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 4, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 5 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 5, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 6 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 6, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte B-Gene Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für B-Gene Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 4, 5 oder 6 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Ein weiterer Gegenstand der Erfindung betrifft daher B-Gene Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 4, 5 oder 6 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

Die Hybridisierungbedingungen sind vorstehend beschrieben.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Aus-

gangsequenz.

- Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.
- Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform B1), B2) oder B3) beschriebenen B-Gene Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 auf.
- Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 als B-Gene Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.
- Alle vorstehend erwähnten B-Gene Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.
- Unter einem „PDS Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Phytoendesaturase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.
- Diese PDS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNABibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen PDS Promotorsequenzen oder den Nukleinsäuren, kodierend eine Phytoende-

saturase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Phytoendesaturase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in
5 der Promotorsequenz.

Unter einer Phytoendesaturase wird vorzugsweise ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluen umzuwandeln.

10. Bevorzugte PDS Promotoren enthalten

- C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder
- C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von
15 Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder
- C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder
- C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3)

20 Die Nukleinsäuresequenz SEQ. ID. NO. 7 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (U46919) dar.

25 Die Nukleinsäuresequenz SEQ. ID. NO. 8 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (X78271) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 9 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (X171023) dar.

30 Die Nukleinsäuresequenz SEQ. ID. NO. 10 stellt eine weitere Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* dar.

35 Die Erfindung betrifft weiterhin PDS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 7, 8, 9 oder 10) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nuklein-säureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist.

40 Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße PDS Promoto- ren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Da-

tenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 7, 8, 9 oder 10 leicht auffinden.

5 Künstliche erfindungsgemäße PDS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

10 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 7 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 7, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

15 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 8 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 8, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 9 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 9, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

25 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 10 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 10, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

30 Besonders bevorzugte PDS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

35 Weitere natürliche Beispiele für PDS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 aus verschiedenen Organismen, deren

genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Ein weiterer Gegenstand der Erfindung betrifft daher PDS Promotoren, enthaltend eine
5 Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

10 Die Hybridisierungbedingungen sind vorstehend beschrieben.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

15 Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

20 Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform C1), C2) oder C3) beschriebenen PDS Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ.
25 ID. NO. 7, 8, 9 oder 10 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 als PDS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

30 Alle vorstehend erwähnten PDS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor
35 Laboratory Press, beschrieben.

40

- Unter einem „CHRC Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend ein Chromoplasten-assoziiertes Protein C regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.
- 5 10 Diese CHRC Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen CHRC Promotorsequenzen oder den Nukleinsäuren kodierend ein Chromoplasten-assoziiertes Protein C auffinden.
- 15 20 Vorzugsweise werden dazu die Nukleinsäuren kodierend ein Chromoplasten-assoziiertes Protein C verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Bevorzugte CHRC Promotoren enthalten

- D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von
25 Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
30 D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3)

Die Nukleinsäuresequenz SEQ. ID. NO. 11 stellt eine Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke (AAV36416) dar.

35 Die Nukleinsäuresequenz SEQ. ID. NO. 12 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 13 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 14 stellt eine weitere Promotorsequenz des chromoplasten assoziiertes Protein C (CHRC) aus Gurke dar.

- Die Erfindung betrifft weiterhin CHRC Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 11, 12, 13, oder 14) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13, oder 14 aufweist.
- 5 Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße CHRC Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 11, 12, 13, oder 14 leicht auffinden.
- 10 Künstliche erfindungsgemäße CHRC Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.
- 15 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 11 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 11, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.
- 20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 12 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.
- 25 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 13 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 13, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.
- 30 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 14 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 14,
- 35

insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

- 5 Besonders bevorzugte CHRC Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

- 10 Weitere natürliche Beispiele für CHRC Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

- 15 Ein weiterer Gegenstand der Erfindung betrifft daher CHCRC Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 11, 12, 13, oder 14 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

- 20 Die Hybridisierungsbedingungen sind vorstehend beschrieben.

- Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

- Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

- Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform D1), D2) oder D3) beschriebenen CHRC Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 als CHRC Promotor, d.h. zur Expression von Genen in Pflanzen der

Gattung Tagetes.

Alle vorstehend erwähnten CHRC Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch

- 5 Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-
10 Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Mit den erfindungsgemäßen Promotoren lässt sich prinzipiell jedes Gen, also jede Nukleinsäure, kodierend ein Protein, in Pflanzen der Gattung Tagetes exprimieren, insbesondere blütenspezifisch exprimieren, besonders bevorzugt petalenspezifisch exprimieren.

- 20 Diese in Pflanzen der Gattung Tagetes zu exprimierenden Gene werden im folgenden auch „Effektgene“ genannt.

Bevorzugte Effektgene sind beispielsweise Gene aus dem Biosyntheseweg von Geruchsstoffen und Blütenfarben, deren Expression oder erhöhte Expression in Pflanzen der Gattung Tagetes zu einer Veränderung der Geruchs und/oder der Blütenfarbe von
25 Blüten der Pflanzen der Gattung Tagetes führt.

Die Biosynthese von flüchtigen Geruchskomponenten, speziell in Blüten, wurde in den letzten Jahren an verschiedenen Modellorganismen wie Clarkia breweri und Antirrhinum majus L. studiert, Flüchtige Geruchskomponenten werden beispielsweise innerhalb des
30 Monoterpen- und Phenylpropan-Stoffwechsels gebildet werden. Im ersten Fall handelt es sich um Linalool; von den Phenylpropanen sind Methyleneugenol, Benzylacetat, Methylbenzoat und Methylsalicat abgeleitet.

Für die Biosynthese von Linalool, (ISo)Methyleugenol, Benzylacetat und Methylsalicinat sind bevorzugte Gene ausgewählt aus der Gruppe Nukleinsäuren kodierend eine Linalool-Synthase (LIS), Nukleinsäuren kodierend eine S-Adenosyl-L-Met:(iso)-Eugenol-O-Methyltransferase (IEMT), Nukleinsäuren kodierend eine Acetyl-CoA-Benzylalkohol-Acetyltransferase und Nukleinsäuren kodierend eine S-Adenosyl-L-Met:Salicylsäure-Methyltransferase (SAMT). Nukleinsäuresequenzen und Proteinsequenzen zu den
40 genannten enzymatischen Aktivitäten sind in Dudareva et al. Plant Cell 8 (1996), 1137-

1148; Wang et al. Plant Physiol. 114 (1997), 213-221 und Dudareva et al. Plant J. 14 (1998) 297-304) beschrieben.

5 Besonders bevorzugte Effektgene sind Gene aus Biosyntheswegen von biosynthetischen Produkten die in Pflanzen der Gattung Tagetes natürlicherweise, d.h. im Wildtyp oder durch genetische Veränderung des Wildtyps hergestellt werden können, insbesondere in Blüten hergestellt werden können, besonders bevorzugt in Petalen hergestellt werden können.

10 Bevorzugte biosynthetische Produkte sind Feinchemikalien.

Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und beinhaltet Verbindungen, die von einem Organismus produziert werden und in verschiedenen Industriezweigen Anwendungen finden, wie bspw., jedoch nicht beschränkt auf die pharmazeutische 15 Industrie, die Landwirtschafts-, Kosmetik, Food und Feed-Industrie. Diese Verbindungen umfassen organische Säuren, wie beispielsweise Weinsäure, Itaconsäure und Diaminopimelinsäure, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Purin- und Pyrimidinbasen, Nukleoside und Nukleotide (wie bspw. beschrieben in Ku- ninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology 20 Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten), Lipide, gesättigte und ungesättigte Fettsäuren (bspw. Arachidonsäure), Diole (bspw. Propan-diol und Butandiol), Kohlenhydrate (bspw. Hyaluronsäure und Trehalose), aromatische Verbindungen (bspw. aromatische Amine, Vanillin und Indigo), Vitamine, Carotinoide und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, 25 Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), Enzyme und 30 sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien). Der Metabolismus und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.

35 I. *Aminosäure-Metabolismus und Verwendungen*

Die Aminosäuren umfassen die grundlegenden Struktureinheiten sämtlicher Proteine und sind somit für die normalen Zellfunktionen essentiell. Der Begriff "Aminosäure" ist im Fachgebiet bekannt. Die proteinogenen Aminosäuren, von denen es 20 Arten gibt, 40 dienen als Struktureinheiten für Proteine, in denen sie über Peptidbindungen miteinan-

der verknüpft sind, wohingegen die nicht-proteinogenen Aminosäuren (von denen Hunderte bekannt sind) gewöhnlich nicht in Proteinen vorkommen (siehe Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97 VCH: Weinheim (1985)). Die Aminosäuren können in der D- oder L-Konfiguration vorliegen, obwohl L-Aminosäuren 5 gewöhnlich der einzige Typ sind, den man in natürlich vorkommenden Proteinen findet. Biosynthese- und Abbauwege von jeder der 20 proteinogenen Aminosäuren sind sowohl bei prokaryotischen als auch eukaryotischen Zellen gut charakterisiert (siehe bspw. Stryer, L. Biochemistry, 3. Auflage, S. 578-590 (1988)). Die "essentiellen" Aminosäuren (Histidin, Isoleucin, Leucin, Lysin, Methionin, Phenylalanin, Threonin, 10 Tryptophan und Valin), so bezeichnet, da sie aufgrund der Komplexität ihrer Biosynthese mit der Ernährung aufgenommen werden müssen, werden durch einfache Biosyntheseswege in die übrigen 11 "nichtessentiellen" Aminosäuren (Alanin, Arginin, Asparagin, Aspartat, Cystein, Glutamat, Glutamin, Glycin, Prolin, Serin und Tyrosin) 15 umgewandelt. Höhere Tiere besitzen die Fähigkeit, einige dieser Aminosäuren zu synthetisieren, jedoch müssen die essentiellen Aminosäuren mit der Nahrung aufgenommen werden, damit eine normale Proteinsynthese stattfindet.

Abgesehen von ihrer Funktion bei der Proteinbiosynthese sind diese Aminosäuren interessante Chemikalien an sich, und man hat entdeckt, daß viele bei verschiedenen Anwendungen in der Nahrungsmittel-, Futter-, Chemie-, Kosmetik-, Landwirtschafts- und pharmazeutischen Industrie zum Einsatz kommen. Lysin ist nicht nur für die Ernährung des Menschen eine wichtige Aminosäure, sondern auch für monogastrische Tiere, wie Geflügel und Schweine. Glutamat wird am häufigsten als Geschmacksadditiv (Mononatriumglutamat, MSG) sowie weithin in der Nahrungsmittelindustrie verwendet, wie 20 auch Aspartat, Phenylalanin, Glycin und Cystein. Glycin, L-Methionin und Tryptophan werden sämtlich in der pharmazeutischen Industrie verwendet. Glutamin, Valin, Leucin, Isoleucin, Histidin, Arginin, Prolin, Serin und Alanin werden in der pharmazeutischen Industrie und der Kosmetikindustrie verwendet. Threonin, Tryptophan und D-/L-Methionin sind weitverbreitete Futtermittelzusätze (Leuchtenberger, W. (1996) Amino 25 acids - technical production and use, S. 466-502 in Rehm et al., (Hrsg.) Biotechnology Bd. 6, Kapitel 14a, VCH: Weinheim). Man hat entdeckt, daß sich diese Aminosäuren außerdem als Vorstufen für die Synthese von synthetischen Aminosäuren und Proteinen, wie N-Acetylcystein, S-Carboxymethyl-L-cystein, (S)-5-Hydroxytryptophan und anderen, in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97, VCH, 30 Weinheim, 1985 beschriebenen Substanzen eignen.

Die Biosynthese dieser natürlichen Aminosäuren in Organismen, die sie produzieren können, bspw. Bakterien, ist gut charakterisiert worden (für einen Überblick der bakteriellen Aminosäure-Biosynthese und ihrer Regulation, s. Umbarger, H.E. (1978) Ann. 40 Rev. Biochem. 47: 533 - 606). Glutamat wird durch reduktive Aminierung von α -

Ketoglutarat, einem Zwischenprodukt im Citronensäure-Zyklus, synthetisiert. Glutamin, Prolin und Arginin werden jeweils nacheinander aus Glutamat erzeugt. Die Biosynthese von Serin erfolgt in einem Dreischritt-Verfahren und beginnt mit 3-Phosphoglycerat (einem Zwischenprodukt bei der Glykolyse), und ergibt nach Oxidations-, Transaminierungs- und Hydrolyseschritten diese Aminosäure. Cystein und Glycin werden jeweils aus Serin produziert, und zwar die erstere durch Kondensation von Homocystein mit Serin, und die letztere durch Übertragung des Seitenketten- β -Kohlenstoffatoms auf Tetrahydrofolat, in einer durch Serintranshydroxymethylase katalysierten Reaktion. Phenylalanin und Tyrosin werden aus den Vorstufen des Glykolyse- und Pentosephosphatweges, Erythrose-4-phosphat und Phosphoenolpyruvat in einem 9-Schritt-Biosyntheseweg synthetisiert, der sich nur in den letzten beiden Schritten nach der Synthese von Prephenat unterscheidet. Tryptophan wird ebenfalls aus diesen beiden Ausgangsmolekülen produziert, jedoch erfolgt dessen Synthese in einem 11-Schritt-Weg. Tyrosin läßt sich in einer durch Phenylalaninhydroxylase katalysierten Reaktion auch aus Phenylalanin herstellen. Alanin; Valin und Leucin sind jeweils Biosyntheseprodukte aus Pyruvat, dem Endprodukt der Glykolyse. Aspartat wird aus Oxalacetat, einem Zwischenprodukt des Citratzyklus, gebildet. Asparagin, Methionin, Threonin und Lysin werden jeweils durch Umwandlung von Aspartat produziert. Isoleucin wird aus Threonin gebildet. In einem komplexen 9-Schritt-Weg erfolgt die Bildung von Histidin aus 5-Phosphoribosyl-1-pyrophosphat, einem aktivierten Zucker.

Aminosäuren, deren Menge den Proteinbiosynthesebedarf der Zelle übersteigt, können nicht gespeichert werden, und werden stattdessen abgebaut, so daß Zwischenprodukte für die Haupt-Stoffwechselwege der Zelle bereitgestellt werden (für einen Überblick siehe Stryer, L., Biochemistry, 3. Aufl. Kap. 21 "Amino Acid Degradation and the Urea Cycle"; S 495-516 (1988)). Die Zelle ist zwar in der Lage, ungewünschte Aminosäuren in nützliche Stoffwechsel-Zwischenprodukte umzuwandeln, jedoch ist die Aminosäureproduktion hinsichtlich der Energie, der Vorstufenmoleküle und der für ihre Synthese nötigen Enzyme aufwendig. Es überrascht daher nicht, daß die Aminosäure-Biosynthese durch Feedback-Hemmung reguliert wird, wobei das Vorliegen einer bestimmten Aminosäure ihre eigene Produktion verlangsamt oder ganz beendet (für einen Überblick über den Rückkopplungs-Mechanismus bei Aminosäure-Biosynthesewegen, siehe Stryer, L., Biochemistry, 3. Aufl., Kap. 24, "Biosynthesis of Amino Acids and Heme", S. 575-600 (1988)). Der Ausstoß einer bestimmten Aminosäure wird daher durch die Menge dieser Aminosäure in der Zelle eingeschränkt.

II. Vitamine, Carotinoide, Cofaktoren und Nutraceutika-Metabolismus sowie Verwendungen

Vitamine, Carotinoide, Cofaktoren und Nutraceutika umfassen eine weitere Gruppe von Molekülen. Höhere Tiere haben die Fähigkeit verloren, diese zu synthetisieren und müssen sie somit aufnehmen, obwohl sie leicht durch andere Organismen, wie Bakterien, synthetisiert werden. Diese Moleküle sind entweder biologisch aktive Moleküle an sich oder Vorstufen von biologisch aktiven Substanzen, die als Elektronenträger oder Zwischenprodukte bei einer Reihe von Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Verarbeitungs-Hilfsstoffe. (Für einen Überblick über die Struktur, Aktivität und die industriellen Anwendungen dieser Verbindungen siehe bspw. Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Der Begriff "Vitamin" ist im Fachgebiet bekannt und umfaßt Nährstoffe, die von einem Organismus für eine normale Funktion benötigt werden, jedoch nicht von diesem Organismus selbst synthetisiert werden können. Die Gruppe der Vitamine kann Cofaktoren und nutraceutische Verbindungen umfassen. Der Begriff "Cofaktor" umfaßt nicht-proteinartige Verbindungen, die für das Auftreten einer normalen Enzymaktivität nötig sind. Diese Verbindungen können organisch oder anorganisch sein; die erfundungsgemäßen Cofaktor-Moleküle sind vorzugsweise organisch. Der Begriff "Nutraceutikum" umfaßt Nahrungsmittelzusätze, die bei Pflanzen und Tieren, insbesondere dem Menschen, gesundheitsfördernd sind. Beispiele solcher Moleküle sind Vitamine, Antioxidantien und ebenfalls bestimmte Lipide (z.B. mehrfach ungesättigte Fettsäuren).

Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Die Biosynthese dieser Moleküle in Organismen, die zu ihrer Produktion befähigt sind, wie Bakterien, ist umfassend charakterisiert worden (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996, Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for free Radical Research - Asien, abgehalten

am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press, Champaign, IL X, 374 S).

- Thiamin (Vitamin B₁) wird durch chemisches Kuppeln von Pyrimidin und Thiazol-Einheiten gebildet. Riboflavin (Vitamin B₂) wird aus Guanosin-5'-triphosphat (GTP) und
5 Ribose-5'-phosphat synthetisiert. Riboflavin wiederum wird zur Synthese von Flavinmononukleotid (FMN) und Flavinadenindinukleotid (FAD) eingesetzt. Die Familie von Verbindungen, die gemeinsam als "Vitamin B6" bezeichnet werden (bspw. Pyridoxin, Pyridoxamin, Pyridoxal-5'-phosphat und das kommerziell verwendete Pyridoxin-hydrochlorid), sind alle Derivate der gemeinsamen Struktureinheit 5-Hydroxy-6-
10 methylpyridin. Panthenat (Pantothenäure, R-(+)-N-(2,4-Dihydroxy-3,3-dimethyl-1-oxobutyl)-β-alanin) kann entweder durch chemische Synthese oder durch Fermentation hergestellt werden. Die letzten Schritte bei der Panthenat-Biosynthese bestehen aus der ATP-getriebenen Kondensation von β-Alanin und Pantoinsäure. Die für die Biosyntheseschritte für die Umwandlung in Pantoinsäure, in β-Alanin und zur Kondensation in
15 Panthenäure verantwortlichen Enzyme sind bekannt. Die metabolisch aktive Form von Panthenat ist Coenzym A, dessen Biosynthese über 5 enzymatische Schritte verläuft. Panthenat, Pyridoxal-5'-phosphat, Cystein und ATP sind die Vorstufen von Coenzym A. Diese Enzyme katalysieren nicht nur die Bildung von Panthenat, sondern auch die Produktion von (R)-Pantoinsäure, (R)-Pantolacton, (R)-Panthenol (Provitamin B₅), Pantethein (und seinen Derivaten) und Coenzym A.

- Die Biosynthese von Biotin aus dem Vorstufenmolekül Pimeloyl-CoA in Mikroorganismen ist ausführlich untersucht worden, und man hat mehrere der beteiligten Gene identifiziert. Es hat sich herausgestellt, daß viele der entsprechenden Proteine an der
25 Fe-Cluster-Synthese beteiligt sind und zu der Klasse der nifS-Proteine gehören. Die Liponsäure wird von der Octanonsäure abgeleitet und dient als Coenzym beim Energie-Metabolismus, wo sie Bestandteil des Pyruvatdehydrogenasekomplexes und des α-Ketoglutaratdehydrogenasekomplexes wird. Die Folate sind eine Gruppe von Substanzen, die alle von der Folsäure abgeleitet werden, die wiederum von L-
30 Glutaminsäure, p-Aminobenzoësäure und 6-Methylpterin hergeleitet ist. Die Biosynthese der Folsäure und ihrer Derivate, ausgehend von den metabolischen Stoffwechselzwischenprodukten Guanosin-5'-triphosphat (GTP), L-Glutaminsäure und p-Aminobenzoësäure ist in bestimmten Mikroorganismen eingehend untersucht worden.
35 Corrinoide (wie die Cobalamine und insbesondere Vitamin B₁₂) und die Porphyrine gehören zu einer Gruppe von Chemikalien, die sich durch ein Tetrapyrrol-Ringsystem auszeichnen. Die Biosynthese von Vitamin B₁₂ ist hinreichend komplex, daß sie noch nicht vollständig charakterisiert worden ist, jedoch ist inzwischen ein Großteil der beteiligten Enzyme und Substrate bekannt. Nikotinsäure (Nikotinat) und Nikotinamid sind
40 Pyridin-Derivate, die auch als "Niacin" bezeichnet werden. Niacin ist die Vorstufe der

wichtigen Coenzyme NAD (Nikotinamidadenindinukleotid) und NADP (Nikotinamidade-nindinukleotidphosphat) und ihrer reduzierten Formen.

Die Produktion dieser Verbindungen im Großmaßstab beruht größtenteils auf zellfreien chemischen Synthesen, obwohl einige dieser Chemikalien ebenfalls durch großangelegte Anzucht von Mikroorganismen produziert worden sind, wie Riboflavin, Vitamin B₆, Pantothenat und Biotin. Nur Vitamin B₁₂ wird aufgrund der Komplexität seiner Synthese lediglich durch Fermentation produziert. In-vitro-Verfahren erfordern einen erheblichen Aufwand an Materialien und Zeit und häufig an hohen Kosten.

10

III. Purin-, Pyrimidin-, Nukleosid- und Nukleotid-Metabolismus und Verwendungen

Gene für den Purin- und Pyrimidin-Stoffwechsel und ihre entsprechenden Proteine sind wichtige Ziele für die Therapie von Tumorerkrankungen und Virusinfektionen. Der Begriff "Purin" oder "Pyrimidin" umfaßt stickstoffhaltige Basen, die Bestandteil der Nukleinsäuren, Coenzyme und Nukleotide sind. Der Begriff "Nukleotid" beinhaltet die grundlegenden Struktureinheiten der Nukleinsäuremoleküle, die eine stickstoffhaltige Base, einen Pentose-Zucker (bei RNA ist der Zucker Ribose, bei DNA ist der Zucker D-Desoxyribose) und Phosphorsäure umfassen. Der Begriff "Nukleosid" umfaßt Moleküle, die als Vorstufen von Nukleotiden dienen, die aber im Gegensatz zu den Nukleotiden keine Phosphorsäureeinheit aufweisen. Durch Hemmen der Biosynthese dieser Moleküle oder ihrer Mobilisation zur Bildung von Nukleinsäuremolekülen ist es möglich, die RNA- und DNA-Synthese zu hemmen; wird diese Aktivität zielgerichtet bei kanzerogenen Zellen gehemmt, läßt sich die Teilungs- und Replikations-Fähigkeit von Tumorzellen hemmen.

Es gibt zudem Nukleotide, die keine Nukleinsäuremoleküle bilden, jedoch als Energiespeicher (d.h. AMP) oder als Coenzyme (d.h. FAD und NAD) dienen.

30 Mehrere Veröffentlichungen haben die Verwendung dieser Chemikalien für diese medizinischen Indikationen beschrieben, wobei der Purin- und/oder Pyrimidin-Metabolismus beeinflußt wird (bspw. Christopherson, R.I. und Lyons, S.D. (1990) "Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents", Med. Res. Reviews 10: 505-548). Untersuchungen an Enzymen, die am Purin- und Pyrimidin-Metabolismus beteiligt sind, haben sich auf die Entwicklung neuer Medikamente konzentriert, die bspw. als Immunsuppressionsmittel oder Antiproliferantien verwendet werden können (Smith, J.L. "Enzymes in Nucleotide Synthesis" Curr. Opin. Struct. Biol. 5 (1995) 752-757; Biochem. Soc. Transact. 23 (1995) 877-902). Die Purin- und Pyrimidinbasen, Nukleoside und Nukleotide haben jedoch auch andere Einsatz-
35 möglichkeiten: als Zwischenprodukte bei der Biosynthese verschiedener Feinchemikalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als
40

- kalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als Energieträger für die Zelle (bspw. ATP oder GTP) und für Chemikalien selbst, werden gewöhnlich als Geschmacksverstärker verwendet (bspw. IMP oder GMP) oder für viele medizinische Anwendungen (siehe bspw. Kuninaka, A., (1996) "Nucleotides and Related Compounds in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim, S. 561-612). Enzyme, die am Purin-, Pyrimidin-, Nukleosid- oder Nukleotid-Metabolismus beteiligt sind, dienen auch immer stärker als Ziele, gegen die Chemikalien für den Pflanzenschutz, einschließlich Fungiziden, Herbiziden und Insektiziden entwickelt werden.
- 10 Der Metabolismus dieser Verbindungen in Bakterien ist charakterisiert worden (für Übersichten siehe bspw. Zalkin, H. und Dixon, J.E. (1992) "De novo purin nucleotide biosynthesis" in Progress in Nucleic Acids Research and Molecular biology, Bd. 42, Academic Press, S. 259-287; und Michal, G. (1999) "Nucleotides and Nucleosides"; Kap. 8 in : Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley, New York). Der Purin-Metabolismus, das Objekt intensiver Forschung, ist für das normale Funktionieren der Zelle essentiell. Ein gestörter Purin-Metabolismus in höheren Tieren kann schwere Erkrankungen verursachen, bspw. Gicht. Die Purinnukleotide werden über eine Reihe von Schritten über die Zwischenverbindung Inosin-5'-phosphat (IMP) aus Ribose-5-phosphat synthetisiert, was zur Produktion von Guanosin-5'-monophosphat (GMP) oder Adenosin-5'-monophosphat (AMP) führt, aus denen sich die als Nukleotide verwendeten Triphosphatformen leicht herstellen lassen. Diese Verbindungen werden auch als Energiespeicher verwendet, so daß ihr Abbau Energie für viele verschiedene biochemische Prozesse in der Zelle liefert. Die Pyrimidinbiosynthese erfolgt über die Bildung von Uridin-5'-monophosphat (UMP) aus Ribose-5-phosphat.
- 15 UMP wiederum wird in Cytidin-5'-triphosphat (CTP) umgewandelt. Die Desoxyformen sämtlicher Nukleotide werden in einer Einschritt-Reduktionsreaktion aus der Diphosphat-Riboseform des Nukleotides zur Diphosphat-Desoxyriboseform des Nukleotides hergestellt. Nach der Phosphorylierung können diese Moleküle an der DNA-Synthese teilnehmen.
- 20
- 25
- 30
- #### IV. Trehalose-Metabolismus und Verwendungen
- Trehalose besteht aus zwei Glucosemolekülen, die über α,α -1,1-Bindung miteinander verknüpft sind. Sie wird gewöhnlich in der Nahrungsmittelindustrie als Süßstoff, als Additiv für getrocknete oder gefrorene Nahrungsmittel sowie in Getränken verwendet. Sie wird jedoch auch in der pharmazeutischen Industrie, der Kosmetik- und Biotechnologie-Industrie angewendet (s. bspw. Nishimoto et al., (1998) US-Patent Nr. 5 759 610; Singer, M.A. und Lindquist, S. Trends Biotech. 16 (1998) 460-467; Paiva, C.L.A. und Panek, A.D. Biotech Ann. Rev. 2 (1996) 293-314; und Shiosaka, M. J. Japan 172 (1997) 97-102). Trehalose wird durch Enzyme von vielen Mikroorganismen produziert

und auf natürliche Weise in das umgebende Medium abgegeben, aus dem sie durch im Fachgebiet bekannte Verfahren gewonnen werden kann.

- 5 Besonders bevorzugte biosynthetische Produkte sind ausgewählt aus der Gruppe organische Säuren, Proteine, Nukleotide und Nukleoside, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Lipide und Fettsäuren, Diole, Kohlenhydrate, aromatische Verbindungen, Vitamine und Cofaktoren, Enzyme und Proteine.

Bevorzugte organische Säuren sind Weinsäure, Itaconsäure und Diaminopimelinsäure

10

Bevorzugte Nukleoside und Nukleotide sind beispielsweise beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten.

15

Bevorzugte biosynthetische Produkte sind weiterhin Lipide, gesättigte und ungesättigte Fettsäuren, wie beispielsweise Arachidonsäure, Diole wie beispielsweise Propandiol und Butandiol, Kohlenhydrate, wie beispielsweise Hyaluronsäure und Trehalose, aromatische Verbindungen, wie beispielsweise aromatische Amine, Vanillin und Indigo, Vitamine und Cofaktoren, wie beispielsweise beschrieben in Ullmann's Encyclopedia of

20

Industrial Chemistry, Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), En-

25

zyme Polyketide (Cane et al. (1998) Science 282: 63-68), und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien.

30

Besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung Tagetes exprimiert werden sind demnach Gene, ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Kohlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein

35

Protein aus dem Biosyntheseweg von Carotinoiden, insbesondere Ketocarotinoiden,

40

Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Cofaktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen.

- Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.
- 10 Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

- Ganz besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung Tagetes exprimiert werden sind demnach Gene die Proteine aus dem Biosyntheseweg von Carotinoiden kodieren.

- Insbesondere bevorzugt sind Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase; Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren 20 kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ε -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase; Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren 25 kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren 30 kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

- Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionen-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

Beispiele für Nukleinsäuren, kodierend eine Ketolase, und die entsprechenden Ketolasen, sind beispielsweise Sequenzen aus

- 5 *Haematoccus pluvialis*, insbesondere aus *Haematoccus pluvialis* Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),
- 10 *Haematoccus pluvialis*, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),
- 15 *Agrobacterium aurantiacum* (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20),
- 20 *Alicaligenes spec.* (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22),
- 25 *Paracoccus marcusii* (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).
- 30 *Synechocystis* sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26).
- 35 *Bradyrhizobium* sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 27, Protein SEQ ID NO: 28).
- 40 *Nostoc* sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 29, Protein SEQ ID NO: 30).
- 45 *Haematococcus pluvialis*
 (Accession NO: AF534876, AAN03484; Nukleinsäure: SEQ ID NO: 31, Protein : SEQ ID NO: 32)
- 50 *Paracoccus* sp. MBIC1143
 (Accession NO: D58420, P54972; Nukleinsäure: SEQ ID NO: 33, Protein : SEQ ID NO: 34)
- 55 *Brevundimonas aurantia*
 (Accession NO: AY166610, AAN86030; Nukleinsäure: SEQ ID NO: 35, Protein : SEQ ID NO: 36)
- 60 *Nodularia spumigena* NSOR10

(Accession NO: AY210783, AAO64399; Nukleinsäure: SEQ ID NO: 37, Protein : SEQ ID NO: 38)

Nostoc punctiforme ATCC 29133

5 (Accession NO: NZ_AABC01000195, ZP_00111258; Nukleinsäure: SEQ ID NO: 39, Protein : SEQ ID NO: 40)

Nostoc punctiforme ATCC 29133

10 (Accession NO: NZ_AABC01000196; Nukleinsäure: SEQ ID NO: 41, Protein : SEQ ID NO: 42)

Deinococcus radiodurans R1

(Accession NO: E75561, AE001872; Nukleinsäure: SEQ ID NO: 43, Protein : SEQ ID NO: 44),

15

Synechococcus sp. WH 8102,

Nukleinsäure: Acc.-No. NZ_AABD01000001, Basenpaar 1,354,725-1,355,528 (SEQ ID NO: 75), Protein: Acc.-No. ZP_00115639 (SEQ ID NO: 76) (als putatives Protein annotiert),

20

Unter einer β -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, einen endständigen, linearen Rest von Lycopin in einen β -Ionon-Ring zu überführen.

25

Insbesondere wird unter einer β -Cyclase ein Protein verstanden, das die enzymatische Aktivität aufweist, γ -Carotin in β -Carotin umzuwandeln.

Beispiele für β -Cyclase-Gene sind Nukleinsäuren, kodierend eine β -Cyclase aus To-

maten (Accession X86452). (Nukleinsäure: SEQ ID NO: 45, Protein: SEQ ID NO: 46),

30

sowie β -Cyclasen der folgenden Accession Nummern:

S66350 lycopene beta-cyclase (EC 5.5.1.-) - tomato

CAA60119 lycopene synthase [Capsicum annuum]

S66349 lycopene beta-cyclase (EC 5.5.1.-) - common tobacco

35

CAA57386 lycopene cyclase [Nicotiana tabacum]

AAM21152 lycopene beta-cyclase [Citrus sinensis]

AAD38049 lycopene cyclase [Citrus x paradisi]

AAQ86060 lycopene cyclase [Citrus unshiu]

AAF44700 lycopene beta-cyclase [Citrus sinensis]

40

AAK07430 lycopene beta-cyclase [Adonis palaestina]

- AAG10429 beta cyclase [*Tagetes erecta*]
AAA81880 lycopene cyclase
AAB53337 Lycopene beta cyclase
AAL92175 beta-lycopene cyclase [*Sandersonia aurantiaca*]
5 CAA67331 lycopene cyclase [*Narcissus pseudonarcissus*]
AAM45381 beta cyclase [*Tagetes erecta*]
AAQ18661 lycopene beta-cyclase [*Zea mays*]
AAG21133 chromoplast-specific lycopene beta-cyclase [*Lycopersicon esculentum*]
AAF18989 lycopene beta-cyclase [*Daucus carota*]
10 ZP_001140 hypothetical protein [*Prochlorococcus marinus str. MIT9313*]
ZP_001050 hypothetical protein [*Prochlorococcus marinus subsp. pastoris str. CCMP1378*]
ZP_001046 hypothetical protein [*Prochlorococcus marinus subsp. pastoris str. CCMP1378*]
15 ZP_001134 hypothetical protein [*Prochlorococcus marinus str. MIT9313*]
ZP_001150 hypothetical protein [*Synechococcus sp. WH 8102*]
AAF10377 lycopene cyclase [*Deinococcus radiodurans*]
BAA29250 393aa long hypothetical protein [*Pyrococcus horikoshii*]
BAC77673 lycopene beta-monocyclase [marine bacterium P99-3]
20 AAL01999 lycopene cyclase [*Xanthobacter sp. Py2*]
ZP_000190 hypothetical protein [*Chloroflexus aurantiacus*]
ZP_000941 hypothetical protein [*Novosphingobium aromaticivorans*]
AAF78200 lycopene cyclase [*Bradyrhizobium sp. ORS278*]
BAB79602 crtY [*Pantoea agglomerans pv. milletiae*]
25 CAA64855 lycopene cyclase [*Streptomyces griseus*]
AAA21262 dycopene cyclase [*Pantoea agglomerans*]
C37802 crtY protein - *Erwinia uredovora*
BAB79602 crtY [*Pantoea agglomerans pv. milletiae*]
AAA64980 lycopene cyclase [*Pantoea agglomerans*]
30 AAC44851 lycopene cyclase
BAA09593 Lycopene cyclase [*Paracoccus sp. MBIC1143*]
ZP_000941 hypothetical protein [*Novosphingobium aromaticivorans*]
CAB56061 lycopene beta-cyclase [*Paracoccus marcusii*]
BAA20275 lycopene cyclase [*Erythrobacter longus*]
35 ZP_000570 hypothetical protein [*Thermobifida fusca*]
ZP_000190 hypothetical protein [*Chloroflexus aurantiacus*]
AAK07430 lycopene beta-cyclase [*Adonis palaestina*]
CAA67331 lycopene cyclase [*Narcissus pseudonarcissus*]
AAB53337 Lycopene beta cyclase
40 BAC77673 lycopene b̄ta-monocyclase [marine bacterium P99-3]

Eine besonders bevorzugte β -Cyclase ist weiterhin die chromoplastenspezifische β -Cyclase aus Tomate (AAG21133) (Nukleinsäure: SEQ. ID. No. 49; Protein: SEQ. ID. No. 50)

5

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

10

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Canthaxanthin in Astaxanthin umzuwandeln.

Beispiele für ein Hydroxylase-Gene sind:

15

eine Nukleinsäure, kodierend eine Hydroxylase aus Haematococcus pluvialis, Accession AX038729, WO 0061764); (Nukleinsäure: SEQ ID NO: 49, Protein: SEQ ID NO: 50),

20

sowie Hydroxylasen der folgenden Accession Nummern:

|emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108_1, AF315289_1, AF296158_1, AAC49443.1, NP_194300.1, NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, AF162276_1, AAO53295.1,

25

AAN85601.1, CRTZ_ERWHE, CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU, CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1, NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1, NP_852012.1, NP_115929.1, ZP_00013255.1

30

Eine besonders bevorzugte Hydroxylase ist weiterhin die Hydroxylase aus Tomate (Accession Y14810, CrtR-b2) (Nukleinsäure: SEQ ID NO: 51; Protein: SEQ ID NO. 52).

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

35

Unter einer (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat in Isopentenyldiphosphat und Dimethylallyldiphosphate umzuwandeln.

40

- Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Hydroxyethyl-ThPP und Glycerinaldehyd-3-Phosphat in 1-Deoxy-D-Xylose-5-Phosphat umzuwandeln.
- 5 Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 1-Deoxy-D-Xylose-5-Phosphat in 2-C-methyl-D-erythritol 4-Phosphat umzuwandeln
- Unter einer Isopentenyl-Diphosphat-Δ-Isomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat in Dimethylallylphosphat umzuwandeln.
- 10 Unter einer Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat und Dimethylallylphosphat in Geranyl-Diphosphat umzuwandeln.
- Unter einer Farnesyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, sequentiell 2 Molekülsopentenyl-Diphosphat mit Dimethylallyl-Diphosphat und dem resultierenden Geranyl-Diphosphat in Farnesyl-Diphosphat umzuwandeln.
- 15 Unter einer Geranyl-Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Farnesyl-Diphosphat und Isopentenyl-Diphosphat in Geranyl-Geranyl-Diphosphat umzuwandeln.
- 20 Unter einer Phytoen-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Geranyl-Geranyl-Diphosphat in Phytoen umzuwandeln.
- Unter einer Phytoen-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluen und/oder Phytofluen in ζ-Carotin (Zetacarotin) umzuwandeln.
- 25 Unter einer Zeta-Carotin-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, ζ-Carotin in Neurosporin und/oder Neurosporin in Lycopin umzuwandeln.
- Unter einem crtISO-Proteins wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 7,9,7',9'-tetra-cis-Lycopin in all-trans-Lycopin umzuwandeln.

Unter einem FtsZ-Protein wird ein Protein verstanden, das eine Zellteilungs und Plastidenteilungs-fördernde Wirkung hat und Homologien zu Tubulinproteinen aufweist.

- Unter einem MinD -Protein wird ein Protein verstanden, das eine multifunktionale Rolle bei der Zellteilung aufweist. Es ist eine Membran-assozierte ATPase und kann innerhalb der Zelle eine oszillierende Bewegung von Pol zu Pol zeigen:

Beispiele für HMG-CoA-Reduktase-Gene sind:

- 10 Eine Nukleinsäure, kodierend eine HMG-CoA-Reduktase aus *Arabidopsis thaliana*, Accession NM_106299; (Nukleinsäure: SEQ ID NO: 53, Protein: SEQ ID NO: 54), sowie weitere HMG-CoA-Reduktase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

15 P54961, P54870, P54868, P54869, O02734, P22791, P54873, P54871, P23228, P13704, P54872, Q01581, P17425, P54874, P54839, P14891, P34135, O64966, P29057, P48019, P48020, P12683, P43256, Q9XEL8, P34136, O64967, P29058, P48022, Q41437, P12684, Q00583, Q9XHL5, Q41438, Q9YAS4, O76819, O28538,

20 Q9Y7D2, P54960, O51628, P48021, Q03163, P00347, P14773, Q12577, Q59468, P04035, O24594, P09610, Q58116, O26662, Q01237, Q01559, Q12649, O74164, O59469, P51639, Q10283, O08424, P20715, P13703, P13702, Q96UG4, Q8SQZ9, O15888, Q9TUM4, P93514, Q39628, P93081, P93080, Q944T9, Q40148, Q84MM0, Q84LS3, Q9Z9N4, Q9KLM0

- 25 Beispiele für (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase-Gene sind:

Eine Nukleinsäure, kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase aus *Arabidopsis thaliana* (lytB/ISPH), ACCESSION 30 AY168881, (Nukleinsäure: SEQ ID NO: 55, Protein: SEQ ID NO:56),

sowie weitere (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

- 35 T04781, AF270978_1, NP_485028.1, NP_442089.1, NP_681832.1, ZP_00110421.1, ZP_00071594.1, ZP_00114706.1, ISPH_SYN3, ZP_00114087.1, ZP_00104269.1, AF398145_1, AF398146_1, AAD55762.1, AF514843_1, NP_622970.1, NP_348471.1, NP_562001.1, NP_223698.1, NP_781941.1, ZP_00080042.1, NP_859669.1, NP_214191.1, ZP_00086191.1, ISPH_VIBCH, NP_230334.1, NP_742768.1,
- 40 NP_302306.1, ISPH_MYCLE, NP_602581.1, ZP_00026966.1, NP_520563.1,

NP_253247.1, NP_282047.1, ZP_00038210.1, ZP_00064913.1, CAA61555.1,
ZP_00125365.1, ISPH_ACICA, EAA24703.1, ZP_00013067.1, ZP_00029164.1,
NP_790656.1, NP_217899.1, NP_641592.1, NP_636532.1, NP_719076.1,
NP_660497.1, NP_422155.1, NP_715446.1, ZP_00090692.1, NP_759496.1,
5 ISPH_BURPS, ZP_00129657.1, NP_215626.1, NP_335584.1, ZP_00135016.1,
NP_789585.1, NP_787770.1, NP_769647.1, ZP_00043336.1, NP_242248.1,
ZP_00008555.1, NP_246603.1, ZP_00030951.1, NP_670994.1, NP_404120.1,
NP_540376.1, NP_733653.1, NP_697503.1, NP_840730.1, NP_274828.1,
NP_796916.1, ZP_00123390.1, NP_824386.1, NP_737689.1, ZP_00021222.1,
10 NP_757521.1, NP_390395.1, ZP_00133322.1, CAD76178.1, NP_600249.1,
NP_454660.1, NP_712601.1, NP_385018.1, NP_751989.1

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene sind:

15 Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase aus Lycopersicon esculentum, ACCESSION #AF143812 (Nukleinsäure: SEQ ID NO: 57 , Protein: SEQ ID NO: 58),
sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene aus anderen Organismen mit den folgenden Accession Nummern:
AF143812_1, DXS_CAPAN, CAD22530.1, AF182286_1, NP_193291.1, T52289,
AAC49368.1, AAP14353.1, D71420, DXS_ORYSA, AF443590_1, BAB02345.1,
CAA09804.2, NP_850620.1, CAD22155.2, AAM65798.1, NP_566686.1, CAD22531.1,
AAC33513.1, CAC08458.1, AAG10432.1, T08140, AAP14354.1, AF428463_1,
25 ZP_00010537.1, NP_769291.1, AAK59424.1, NP_107784.1, NP_697464.1,
NP_540415.1, NP_196699.1, NP_384986.1, ZP_00096461.1, ZP_00013656.1,
NP_353769.1, BAA83576.1, ZP_00005919.1, ZP_00006273.1, NP_420871.1,
AAM48660.1, DXS_RHOCA, ZP_00045608.1, ZP_00031686.1, NP_841218.1,
ZP_00022174.1, ZP_00086851.1, NP_742690.1, NP_520342.1, ZP_00082120.1,
30 NP_790545.1, ZP_00125266.1, CAC17468.1, NP_252733.1, ZP_00092466.1,
NP_439591.1, NP_414954.1, NP_752465.1, NP_622918.1, NP_286162.1,
NP_836085.1, NP_706308.1, ZP_00081148.1, NP_797065.1, NP_213598.1,
NP_245469.1, ZP_00075029.1, NP_455016.1, NP_230536.1, NP_459417.1,
NP_274863.1, NP_283402.1, NP_759318.1, NP_406652.1, DXS_SYNLE,
35 DXS_SYNP7, NP_440409.1, ZP_00067331.1, ZP_00122853.1, NP_717142.1,
ZP_00104889.1, NP_243645.1, NP_681412.1, DXS_SYNEL, NP_637787.1,
DXS_CHLTE, ZP_00129863.1, NP_661241.1, DXS_XANCP, NP_470738.1,
NP_484643.1, ZP_00108360.1, NP_833890.1, NP_846629.1, NP_658213.1,
NP_642879.1, ZP_00039479.1, ZP_00060584.1, ZP_00041364.1, ZP_00117779.1,

NP_299528.1

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene sind:

- 5 Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase aus *Arabidopsis thaliana*, ACCESSION #AF148852, (Nukleinsäure: SEQ ID NO: 59, Protein: SEQ ID NO: 60),
10 sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

AF148852, AY084775, AY054682, AY050802, AY045634, AY081453, AY091405, AY098952; AJ242588, AB009053, AY202991, NP_201085.1, T52570, AF331705_1, BAB16915.1, AF367205_1, AF250235_1, CAC03581.1, CAD22156.1, AF182287_1,
15 DXR_MENPI, ZP_00071219.1, NP_488391.1, ZP_00111307.1, DXR_SYNLE, AAP56260.1, NP_681831.1, NP_442113.1, ZP_00115071.1, ZP_00105106.1, ZP_00113484.1, NP_833540.1, NP_657789.1, NP_661031.1, DXR_BACHD, NP_833080.1, NP_845693.1, NP_562610.1, NP_623020.1, NP_810915.1, NP_243287.1, ZP_00118743.1, NP_464842.1; NP_470690.1, ZP_00082201.1,
20 NP_781898.1, ZP_00123667.1, NP_348420.1, NP_604221.1, ZP_00053349.1, ZP_00064941.1, NP_246927.1, NP_389537.1, ZP_00102576.1, NP_519531.1, AF124757_19, DXR_ZYMMO, NP_713472.1, NP_459225.1, NP_454827.1, ZP_00045738.1, NP_743754.1, DXR_PSEPK, ZP_00130352.1, NP_702530.1, NP_841744.1, NP_438967.1, AF514841_1, NP_706118.1, ZP_00125845.1,
25 NP_404661.1, NP_285867.1, NP_240064.1, NP_414715.1, ZP_00094058.1, NP_791365.1, ZP_00012448.1, ZP_00015132.1, ZP_00091545.1, NP_629822.1, NP_771495.1, NP_798691.1, NP_231885.1, NP_252340.1, ZP_00022353.1, NP_355549.1, NP_420724.1, ZP_00085169.1, EAA17616.1, NP_273242.1, NP_219574.1, NP_387094.1, NP_296721.1, ZP_00004209.1, NP_823739.1,
30 NP_282934.1, BAA77848.1; NP_660577.1, NP_760741.1, NP_641750.1, NP_636741.1, NP_829309.1, NP_298338.1, NP_444964.1, NP_717246.1, NP_224545.1, ZP_00038451.1, DXR_KITGR, NP_778563.1.

Beispiele für Isopentenyl-Diphosphat-Δ-Isomerase-Gene sind:

35 Eine Nukleinsäure, kodierend eine Isopentenyl-Diphosphat-Δ-Isomerase aus *Adonis palaestina* clone ApIPI28, (ipiAa1), ACCESSION #AF188060, veröffentlicht durch Cunningham,F.X. Jr. and Gantt,E.: Identification of multi-gene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli*, Plant Cell Physiol. 41 (1), 119-123 (2000) (Nukleinsäure: SEQ ID NO: 61; Protein:
40

SEQ ID NO: 62),

sowie weitere Isopentenyl-Diphosphat-Δ-Isomerase–Gene aus anderen Organismen mit den folgenden Accession Nummern:

5

Q38929, O48964, Q39472, Q13907, O35586, P58044, O42641, O35760, Q10132, P15496, Q9YB30, Q8YNH4, Q42553, O27997, P50740, O51627, O48965, Q8KFR5, Q39471, Q39664, Q9RVE2, Q01335, Q9HHE4, Q9BXS1, Q9KWF6, Q9CIF5, Q88WB6, Q92BX2, Q8Y7A5, Q8TT35, Q9KK75, Q8NN99, Q8XD58, Q8FE75,

10

Q46822, Q9HP40, P72002, P26173, Q9Z5D3, Q8Z3X9, Q8ZM82, Q9X7Q6, O13504, Q9HFW8, Q8NJL9, Q9UUQ1, Q9NH02, Q9M6K9, Q9M6K5, Q9FXR6, O81691, Q9S7C4, Q8S3L8, Q9M592, Q9M6K3, Q9M6K7, Q9FV48, Q9LLB6, Q9AVJ1, Q9AVG8, Q9M6K6, Q9AVJ5, Q9M6K2, Q9AYS5, Q9M6K8, Q9AVG7, Q8S3L7, Q8W250, Q94IE1, Q9AVI8, Q9AYS6, Q9SAY0, Q9M6K4, Q8GVZ0, Q84RZ8,

15

Q8KZ12, Q8KZ66, Q8FND7, Q88QC9, Q8BFZ6, BAC26382, CAD94476.

Beispiele für Geranyl-Diphosphat-Synthase -Gene sind:

Eine Nukleinsäure, kodierend eine Geranyl-Diphosphat-Synthase aus *Arabidopsis thali-*

20

iana, ACCESSION #Y17376, Bouvier,F., Suire,C., d'Harlingue,A., Backhaus,R.A. and Camara,B.; Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpane synthesis in plant cells, Plant J. 24 (2), 241-252 (2000) (Nukleinsäure: SEQ ID NO: 63, Protein: SEQ ID NO: 64),

25

sowie weitere Geranyl-Diphosphat-Synthase–Gene aus anderen Organismen mit den folgenden Accession Nummern:

Q9FT89, Q8LKJ2, Q9FSW8, Q8LKJ3, Q9SBR3, Q9SBR4, Q9FET8, Q8LKJ1, Q84LG1, Q9JK86

30

Beispiele für Farnesyl-Diphosphat-Synthase-Gene sind:

Eine Nukleinsäure, kodierend eine Farnesyl-Diphosphat-Synthase aus *Arabidopsis thaliana* (FPS1), ACCESSION #U80605, veröffentlicht durch Cunillera,N., Arro,M., Deloume,D., Karst,F., Boronat,A. und Ferrer,A.: *Arabidopsis thaliana* contains two differentially expressed farnesyl-diphosphate synthase genes, J. Biol. Chem. 271 (13), 7774-7780 (1996), (Nukleinsäure: SEQ ID NO: 65, Protein: SEQ ID NO:66),

sowie weitere Farnesyl-Diphosphat-Synthase–Gene aus anderen Organismen mit den folgenden Accession Nummern:

P53799, P37268, Q02769, Q09152, P49351, O24241, Q43315, P49352, O24242,
5 P49350, P08836, P14324, P49349, P08524, O66952, Q08291, P54383, Q45220,
P57537, Q8K9A0, P22939, P45204, O66126, P55539, Q9SWH9, Q9AVI7, Q9FRX2,
Q9AYS7, Q94IE8, Q9FXR9, Q9ZWF6, Q9FXR8, Q9AR37, O50009, Q94IE9, Q8RVK7,
Q8RVQ7, O04882, Q93RA8, Q93RB0, Q93RB4, Q93RB5, Q93RB3, Q93RB1,
Q93RB2, Q920E5.

10

Beispiele für Geranyl-geranyl-Diphosphat-Synthase -Gene sind:

Eine Nukleinsäure, kodierend eine Geranyl-geranyl-Diphosphat-Synthase aus Sinaps alba, ACCESSION #X98795, veröffentlicht durch Bonk,M., Hoffmann,B., Von Lintig,J.,

15 Schledz,M., Al-Babili,S., Hobeika,E., Kleinig,H. and Beyer,P.: Chloroplast import of four carotenoid biosynthetic enzymes in vitro reveals differential fates prior to membrane binding and oligomeric assembly, Eur. J. Biochem. 247 (3), 942-950 (1997), (Nuklein-säure: SEQ ID NO: 67, Protein: SEQ ID NO: 68),

20 sowie weitere Geranyl-geranyl-Diphosphat-Synthase–Gene aus anderen Organismen mit den folgenden Accession Nummern:

P22873, P34802 ,P56966, P80042, Q42698, Q92236, O95749, Q9WTN0, Q50727,
P24322, P39464, Q9FXR3, Q9AYN2, Q9FXR2, Q9AVG6, Q9FRW4, Q9SXZ5,
25 Q9AVJ7, Q9AYN1, Q9AVJ4, Q9FXR7, Q8LSC5, Q9AVJ6, Q8LSC4, Q9AVJ3,
Q9SSU0, Q9SXZ6, Q9SST9, Q9AVJ0, Q9AVI9, Q9FRW3, Q9FXR5, Q94IF0,
Q9FRX1, Q9K567, Q93RA9, Q93QX8, CAD95619, EAA31459

Beispiele für Phytoen-Synthase-Gene sind:

30

Eine Nukleinsäure, kodierend eine Phytoen-Synthase aus Erwinia uredovora, ACCES-
SION # D90087; veröffentlicht durch Misawa,N., Nakagawa,M., Kobayashi,K., Yama-
no,S., Izawa,Y., Nakamura,K. und Harashima,K.: Elucidation of the Erwinia uredovora
carotenoid biosynthetic pathway by functional analysis of gene products expressed in

35 Escherichia coli; J. Bacteriol. 172 (12), 6704-6712 (1990), (Nukleinsäure: SEQ ID NO:
69, Protein: SEQ ID NO: 70),

sowie weitere Phytoen-Synthase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

40

CAB39693, BAC69364, AAF10440, CAA45350, BAA20384, AAM72615, BAC09112,
CAA48922, P_001091, CAB84588, AAF41518, CAA48155, AAD38051, AAF33237,
AAG10427, AAA34187, BAB73532, CAC19567, AAM62787, CAA55391, AAB65697,
AAM45379, CAC27383, AAA32836, AAK07735, BAA84763, P_000205, AAB60314,
5 P_001163, P_000718, AAB71428, AAA34153, AAK07734, CAA42969, CAD76176,
CAA68575, P_000130, P_001142, CAA47625, CAA85775, BAC14416, CAA79957,
BAC76563, P_000242, P_000551, AAL02001, AAK15621, CAB94795, AAA91951,
P_000448

10 Beispiele für Phytoen-Desaturase-Gene sind:

Eine Nukleinsäure, kodierend eine Phytoen-Desaturase aus Erwinia uredovora, AC-
CESSION # D90087; veröffentlicht durch Misawa,N., Nakagawa,M., Kobayashi,K.,
Yamano,S., Izawa,Y., Nakamura,K. und Harashima,K.: Elucidation of the Erwinia ure-
15 dovora carotenoid biosynthetic pathway by functional analysis of gene products ex-
pressed in Escherichia coli; J. Bacteriol. 172 (12), 6704-6712 (1990), (Nukleinsäure:
SEQ ID NO: 71, Protein: SEQ ID NO: 72),

sowie weitere Phytoen-Desaturase -Gene aus anderen Organismen mit den folgenden
20 Accession Nummern:

AAL15300, A39597, CAA42573, AAK51545, BAB08179, CAA48195, BAB82461,
AAK92625, CAA55392, AAG10426, AAD02489, AAO24235, AAC12846, AAA99519,
AAL38046, CAA60479, CAA75094, ZP_001041, ZP_001163, CAA39004, CAA44452,
25 ZP_001142, ZP_000718, BAB82462, AAM45380, CAB56040, ZP_001091, BAC09113,
AAP79175, AAL80005, AAM72642, AAM72043, ZP_000745, ZP_001141, BAC07889,
CAD55814, ZP_001041, CAD27442, CAE00192, ZP_001163, ZP_000197, BAA18400,
AAG10425, ZP_001119, AAF13698, 2121278A, AAB35386, AAD02462, BAB68552,
CAC85667, AAK51557, CAA12062, AAG51402, AAM63349, AAF85796, BAB74081,
30 AAA91161, CAB56041, AAC48983, AAG14399, CAB65434, BAB73487, ZP_001117,
ZP_000448, CAB39695, CAD76175, BAC69363, BAA17934, ZP_000171, AAF65586,
ZP_000748, BAC07074, ZP_001133, CAA64853, BAB74484, ZP_001156, AAF23289,
AAG28703, AAP09348, AAM71569, BAB69140, ZP_000130, AAF41516, AAG18866,
35 CAD95940, NP_656310, AAG10645, ZP_000276, ZP_000192, ZP_000186,
AAM94364, EAA31371, ZP_000612, BAC75676, AAF65582

Beispiele für Zeta-Carotin-Desaturase-Gene sind:

Eine Nukleinsäure, kodierend eine Zeta-Carotin-Desaturase aus Narcissus pseudonar-
40 cissus, ACCESSION #AJ224683, veröffentlicht durch Al-Babili,S., Oelschlegel,J. and

Beyer,P.: A cDNA encoding for beta carotene desaturase (Accession No.AJ224683) from Narcissus pseudonarcissus L.. (PGR98-103), Plant Physiol. 117, 719-719 (1998), (Nukleinsäure: SEQ ID NO: 73, Protein: SEQ ID NO: 74),

5 sowie weitere Zeta-Carotin-Desaturase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

Q9R6X4, Q38893, Q9SMJ3, Q9SE20, Q9ZTP4, O49901, P74306, Q9FV46, Q9RCT2,
10 ZDS_NARPS, BAB68552.1, CAC85667.1, AF372617_1, ZDS_TARER, CAD55814.1,
CAD27442.1, 2121278A, ZDS_CAPAN, ZDS_LYCES, NP_187138.1, AAM63349.1,
ZDS_ARATH, AAA91161.1, ZDS_MAIZE, AAG14399.1, NP_441720.1, NP_486422.1,
ZP_00111920.1, CAB56041.1, ZP_00074512.1, ZP_00116357.1, NP_681127.1,
ZP_00114185.1, ZP_00104126.1, CAB65434.1, NP_662300.1

15 Beispiele für crtISO-Gene sind:

Eine Nukleinsäure, kodierend eine crtISO aus Lycopersicon esculentum; ACCESSION #AF416727, veröffentlicht durch Isaacson,T., Ronen,G., Zamir,D. and Hirschberg,J.: Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants; Plant Cell 14 (2), 333-342 (2002), (Nukleinsäure: SEQ ID NO: 75, Protein: SEQ ID NO: 76),

sowie weitere crtISO -Gene aus anderen Organismen mit den folgenden Accession Nummern:

25
AAM53952

Beispiele für FtsZ-Gene sind:

30 Eine Nukleinsäure, kodierend eine FtsZ aus Tagetes erecta, ACCESSION #AF251346, veröffentlicht durch Moehs,C.P., Tian,L., Osteryoung,K.W. and Dellapenna,D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development Plant Mol. Biol. 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 77, Protein: SEQ ID NO: 78),

35 sowie weitere FtsZ -Gene aus anderen Organismen mit den folgenden Accession Nummern:

CAB89286.1, AF205858_1, NP_200339.1, CAB89287.1, CAB41987.1, AAA82068.1,
40 T06774,AF383876_1, BAC57986.1, CAD22047.1, BAB91150.1, ZP_00072546.1,

NP_440816.1, T51092, NP_683172.1, BAA85116.1, NP_487898.1, JC4289,
BAA82871.1, NP_781763.1, BAC57987.1, ZP_00111461.1, T51088, NP_190843.1,
ZP_00060035.1, NP_846285.1, AAL07180.1, NP_243424.1, NP_833626.1,
AAN04561.1, AAN04557.1, CAD22048.1, T51089, NP_692394.1, NP_623237.1,
5 NP_565839.1, T51090, CAA07676.1, NP_113397.1, T51087, CAC44257.1, E84778,
ZP_00105267.1, BAA82091.1, ZP_00112790.1, BAA96782.1, NP_348319.1,
NP_471472.1, ZP_00115870.1, NP_465556.1, NP_389412.1, BAA82090.1,
NP_562681.1, AAM22891.1, NP_371710.1, NP_764416.1, CAB95028.1,
FTSZ_STRGR, AF120117_1, NP_827300.1, JE0282, NP_626341.1, AAC45639.1,
10 NP_785689.1, NP_336679.1, NP_738660.1, ZP_00057764.1, AAC32265.1,
NP_814733.1, FTSZ_MYCKA, NP_216666.1, CAA75616.1, NP_301700.1,
NP_601357.1, ZP_00046269.1, CAA70158.1, ZP_00037834.1, NP_268026.1,
FTSZ_ENTHR, NP_787643.1, NP_346105.1, AAC32264.1, JC5548, AAC95440.1,
NP_710793.1, NP_687509.1, NP_269594.1, AAC32266.1, NP_720988.1,
15 NP_657875.1, ZP_00094865.1, ZP_00080499.1, ZP_00043589.1, JC7087,
NP_660559.1, AAC46069.1, AF179611_14, AAC44223.1, NP_404201.1.

Beispiele für MinD -Gene sind:

20 Eine Nukleinsäure, kodierend eine MinD aus Tagetes erecta, ACCESSION
#AF251019, veröffentlicht durch Moehs,C.P., Tian,L., Osteryoung,K.W. und Dellapenna,D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development; Plant Mol. Biol. 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 79, Protein: SEQ ID NO: 80),
25 sowie weitere MinD –Gene mit den folgenden Accession Nummern:

NP_197790.1, BAA90628.1, NP_038435.1, NP_045875.1, AAN33031.1,
NP_050910.1, CAB53105.1, NP_050687.1, NP_682807.1, NP_487496.1,
30 ZP_00111708.1, ZP_00071109.1, NP_442592.1, NP_603083.1, NP_782631.1,
ZP_00097367.1, ZP_00104319.1, NP_294476.1, NP_622555.1, NP_563054.1,
NP_347881.1, ZP_00113908.1, NP_834154.1, NP_658480.1, ZP_00059858.1,
NP_470915.1, NP_243893.1, NP_465069.1, ZP_00116155.1, NP_390677.1,
NP_692970.1, NP_298610.1, NP_207129.1, ZP_00038874.1, NP_778791.1,
35 NP_223033.1, NP_641561.1, NP_636499.1, ZP_00088714.1, NP_213595.1,
NP_743889.1, NP_231594.1, ZP_00085067.1, NP_797252.1, ZP_00136593.1,
NP_251934.1, NP_405629.1, NP_759144.1, ZP_00102939.1, NP_793645.1,
NP_699517.1, NP_460771.1, NP_860754.1, NP_456322.1, NP_718163.1,
NP_229666.1, NP_357356.1, NP_541904.1, NP_287414.1, NP_660660.1,
40 ZP_00128273.1, NP_103411.1, NP_785789.1, NP_715361.1, AF149810_1,

NP_841854.1, NP_437893.1, ZP_00022726.1, EAA24844.1, ZP_00029547.1,
NP_521484.1, NP_240148.1, NP_770852.1, AF345908_2, NP_777923.1,
ZP_00048879.1, NP_579340.1, NP_143455.1, NP_126254.1, NP_142573.1,
NP_613505.1, NP_127112.1, NP_712786.1, NP_578214.1, NP_069530.1,
5 NP_247526.1, AAA85593.1, NP_212403.1, NP_782258.1, ZP_00058694.1,
NP_247137.1, NP_219149.1, NP_276946.1, NP_614522.1, ZP_00019288.1,
CAD78330.1

Die Erfindung betrifft ferner eine genetisch veränderte Pflanze der Gattung Tagetes,
10 wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch die erfindungsgemäß Promotoren.

15 Wie vorstehend erwähnt, wird unter „Expressionsaktivität“ erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate verstanden.

Unter „spezifischer Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten
20 Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

Bei einer „verursachten Expressionsaktivität“ oder „verursachten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp der Pflanze der Gattung Tagetes
25 nicht vorhanden war.

Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes kein Ketolase-Gen auf. Die Regulation der Expression des Ketolase-Gens in der Pflanze durch die erfindungsgemäß Promotoren führt somit zu einer Verursachung der Expressionsrate.

30 Bei einer „erhöhten Expressionsaktivität“ oder „erhöhten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp der Pflanze der Gattung Tagetes in einer bestimmten Zeit die gebildete Menge des Proteins erhöht.

35 Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes ein Hydroxylase-Gen auf. Die Regulation der Expression des Hydroxylase-Gens in der Pflanze durch die erfindungsgemäß Promotoren führt somit zu einer Erhöhung der Expressionsrate.

40 In einer bevorzugten Ausführungsform der erfindungsgemäß genetisch veränderten Pflanzen der Gattung Tagetes wird die Regulation der Expression von Genen in der

Pflanze durch die erfindungsgemäßen Promotoren dadurch erreicht, dass man

- 5 a) einen oder mehrere erfindungsgemäße Promotoren in das Genom der Pflanze ein-bringt, so dass die Expression eines oder mehrerer endogenen Gene unter der Kon-trolle der eingebrachten erfindungsgemäßen Promotoren erfolgt oder
- b) ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen, erfin-dungsgemäßen Promotoren erfolgt oder

10

- c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungs-gemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Ge-ne in die Pflanze einbringt.

15

- In einer bevorzugten Ausführungsform bringt man gemäß Merkmal c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungsgemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Gene, in die Pflanze ein. Die Integration der Nukleinsäurekonstrukte in der Pflanze der Gattung Tagetes kann dabei intrachromosomal oder extrachromosomal erfolgen.

20

Bevorzugte erfindungsgemäße Promotoren und bevorzugte zu exprimierende Gene (Effektgene) sind vorstehend beschrieben.

25

Im folgenden wird exemplarisch die Herstellung der genetisch veränderten Pflanzen der Gattung Tagetes mit erhöhter oder verursachter Expressionsrate eines Effektgens beschrieben.

Die Transformation kann bei den Kombinationen von genetischen Veränderungen ein-zeln oder durch Mehrfachkonstrukte erfolgen.

30

Die Herstellung der transgenen Pflanzen erfolgt vorzugsweise durch Transformation der Ausgangspflanzen, mit einem Nukleinsäurekonstrukt, das mindestens einen der vorstehend beschriebenen, erfindungsgemäßen Promotoren enthält, die mit einem zu exprimierenden Effektgen und gegebenenfalls weiteren Regulationssignalen funktionell verknüpft sind.

35

Diese Nukleinsäurekonstrukte, in denen die erfindungsgemäßen Promotoren und Ef-fektgene funktionell verknüpft sind, werden im folgenden auch Expressionskassetten genannt.

- Die Expressionskassetten können weitere Regulationssignale enthalten, also regulative Nukleinsäuresequenzen, welche die Expression der Effektgene in der Wirtszelle steuern. Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der kodierenden Sequenz, mindestens einen erfindungsgemäßen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungs-signal und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwi-schenliegenden kodierenden Sequenz des Effektgens für mindestens eines der vor-stehend beschriebenen Gene operativ verknüpft sind.
- 5 Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, das jedes der regulativen Elemente seine Funktion bei der Expression der ko-dierenden Sequenz bestimmungsgemäß erfüllen kann.
- 10 Im folgenden werden beispielhaft die bevorzugten Nukleinsäurekonstrukte, Expressi-onskassetten und Vektoren für Pflanzen und Verfahren zur Herstellung von transgenen Pflanzen, sowie die transgenen Pflanzen der Gattung Tagetes selbst beschrieben.
- 15 Die zur operativen Verknüpfung bevorzugten, aber nicht darauf beschränkten Sequen-zien, sind Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten und Translationsverstärkern wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693 -8711).
- 20
- 25 Die Herstellung einer Expressionskassette erfolgt vorzugsweise durch Fusion mindes-tens eines erfindungsgemäßen Promotors mit mindestens einem Gen, vorzugsweise mit einem der vorstehend beschriebenen Effektgene, und vorzugsweise einer zwischen Promotor und Nukleinsäure-Sequenz inserierten Nukleinsäure, die für ein plastiden-spezifisches Transitpeptid kodiert, sowie einem Polyadenylierungssignal nach gängi-gen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987), beschrieben sind.
- 30
- 35

Die vorzugsweise insertierte Nukleinsäuren, kodierend ein plastidäres Transitpeptid, gewährleisten die Lokalisation in Plastiden und insbesondere in Chromoplasten.

Es können auch Expressionskassetten verwendet werden, deren Nukleinsäure—

- 5 Sequenz für ein Effektgen-Produkt-Fusionsprotein kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das die Translokation des Polypeptides steuert. Bevorzugt sind für die Chromoplasten spezifische Transitpeptide, welche nach Translokation der Effektgene in die Chromoplasten vom Effektgenprodukt-Teil enzymatisch abgespalten werden.

10

Insbesondere bevorzugt ist das Transitpeptid, das von der plastidären *Nicotiana tabacum* Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco (*rbcS*) oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2 oder dessen funktionellem Äquivalent

15 abgeleitet ist.

Besonders bevorzugt sind Nukleinsäure-Sequenzen von drei Kassetten des Plastiden-Transitpeptids der plastidären Transketolase aus Tabak in drei Leserastern als KpnI/BamHI Fragmente mit einem ATG-Codon in der NcoI Schnittstelle:

20

pTP09

KpnI_GGTACCATGGCGTCTTCTTCTCTCACTCTCTCAAGCTATCCTCTC
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTCAACTTTCCCCTTCTCT-

25 CACTTTTCGGCCTAAATCCAATCCAATATCACCCACCTCCGCCGCG-
TACTCCTTCCTCCGCCGCCGCCGCCGCGTAAGGTACCGGC-
GATTCTGTGCCTCAGCTGCAACCGAAACCATAAGAGAAA
GATCC_BamHI

30 pTP10

KpnI_GGTACCATGGCGTCTTCTTCTCTCACTCTCTCAAGCTATCCTCTC
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTCAACTTTCCCCTTCTCT-

35 CACTTTTCGGCCTAAATCCAATCCAATATCACCCACCTCCGCCGCG-
TACTCCTTCCTCCGCCGCCGCCGCCGCGTAAGGTACCGGC-
GATTCTGTGCCTCAGCTGCAACCGAAACCATAAGAGAAA
GATCC_BamHI

pTP11

40

KpnI_GGTACCATGGCGTCTTCTTCTCACTCTCTCAAGCTATCCTCTCT
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTCAACTTCCCCTTCTCT-
CACTTTCCGGCCTAAATCCAATCCAATTCACCAACCTCCGCCGCCG-
TACTCCTCCTCCGCCGCCGCCGTCGTAAGGTACCCGGC-
5 GATT CGT GCCT CAG CTG CA ACC GAA ACC ATAG AG AAA ACT GAG ACT GC GGG-
GATCC_BamHI

- Weitere Beispiele für ein plastidäres Transitpeptid sind das Transitpeptid der plastidären Isopentenyl-pyrophosphat Isomerase-2 (IPP-2) aus *Arabisopsis thaliana* und das
10 Transitpeptid der kleinen Untereinheit der Ribulosebisphosphat Carboxylase (*rbcS*) aus Erbse (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. *Nucl. Acids Res.* 16: 11380).
- Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich
15 gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure- Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen.
- Bevorzugt sind, wie vorstehend beschrieben, synthetische Nukleotid-Sequenzen mit
20 Kodons, die von Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können aus Kodons mit der höchsten Proteinhäufigkeit bestimmt werden, die in den meisten interessanten Pflanzenspezies exprimiert werden.
- Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente
25 manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adapto-
ren oder Linker angesetzt werden.
- 30 Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allge-
meinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger
35 als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zur Wirtspflanze sein. Die Expressionskassette beinhaltet vorzugsweise in der 5'-3'-Transkriptionsrichtung den Promotor, eine kodierende Nukleinsäuresequenz oder ein Nukleinsäurekonstrukt und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche

sind gegeneinander beliebig austauschbar.

Beispiele für einen Terminator sind der 35S-Terminator (Guerineau et al. (1988) Nucl Acids Res. 16: 11380), der nos Terminator (Depicker A, Stachel S; Dhaese P,

- 5 Zambryski P, Goodman HM. Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1982;1(6):561-73) oder der ocs Terminator (Gielen, J, de Beuckeleer, M, Seurinck, J, Debroek, H, de Greve, H, Lemmers, M, van Montagu, M, Schell, J (1984) The complete sequence of the T-DNA of the Agrobacterium tumefaciens plasmid pTiAch5. EMBO J. 3: 835-846).

10

Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt werden. Wo Insertionen, Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können *in vitro*-Mutagenese, "primer-repair", Restriktion oder Ligation verwendet werden.

Bei geeigneten Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "bluntends", können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

20

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus Agrobacterium tumefaciens, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente.

Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet.

30

Dazu können an sich bekannte Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt werden.

35

Geeignete Methoden zur Transformation von Pflanzen sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone – die sogenannte "particle bombardment" Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der, vorstehend beschriebene, durch *Agrobacterium* vermittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene

40

Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von

S.D. Kung und R. Wu, Academic Press (1993), 128-143 sowie in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225) beschrieben.

5 Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, *Agrobacterium tumefaciens* zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) oder besonders bevorzugt pSUN2, pSUN3, pSUN4 oder pSUN5 (WO 02/00900).

10 Mit einem Expressionsplasmid transformierte Agrobakterien können in bekannter Weise zur Transformation von Pflanzen verwendet werden, z.B. indem verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

15 Zur bevorzugten Herstellung von genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, wird die fusionierte Expressionskassette in einen Vektor, beispielsweise pBin19 oder insbesondere pSUN5 und pSUN3 kloniert, der geeignet ist, in *Agrobacterium tumefaciens* transformiert zu werden. Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

20 Die Transformation von Pflanzen durch Agrobakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die ein oder mehrere in die Expressionskassette integrierte Gene enthalten.

25 30 Zur Transformation einer Wirtspflanze mit einem oder mehreren erfindungsgemäßen Effektgenen wird eine Expressionskassette als Insertion in einen rekombinanten Vektor eingebaut, dessen Vektor-DNA zusätzliche funktionelle Regulationssignale, beispielsweise Sequenzen für Replikation oder Integration enthält. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrieben.

35 Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Expressionskassetten in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a. pJIT117 (Guerineau et al. (1988) Nucl. Acids Res. 16 :11380), pBR332, pUC-Serien,

M13mp-Serien und pACYC184. Besonders geeignet sind binäre Vektoren, die sowohl in *E. coli* als auch in Agrobakterien replizieren können.

- Die Erfindung betrifft daher weiterhin eine genetisch veränderte Pflanze der Gattung Tagetes, enthaltend einen erfindungsgemäßen Promotor und funktionell verknüpft ein zu eprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
- 10 Bevorzugte erfindungsgemäße Promotoren und bevorzugte Effektgene sind vorstehend beschrieben.

Insbesondere bevorzugt sind Effektgene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ϵ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Bevorzugte, genetisch veränderte Pflanzen der Gattung Tagetes sind Marigold, Tagetes erecta, Tagetes patula, Tagetes lucida, Tagetes pringlei, Tagetes palmeri, Tagetes minuta oder Tagetes campanulata.

Durch die erfindungsgemäßen Promotoren ist es mit Hilfe der vorstehend beschriebenen, erfindungsgemäßen Verfahren möglich, in den vorstehend beschriebenen, erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes die Stoffwechselwege zu spezifischen biosynthetischen Produkten zu regulieren.

Dazu werden beispielsweise Stoffwechselwege, die zu einem spezifischen biosynthetischen Produkt führen, durch Verursachung oder Erhöhung der Transkriptionrate bzw. Expressionsrate von Genen dieses Biosyntheseweges verstärkt, indem die erhöhte

Proteinmenge zu einer erhöhten Gesamtaktivität dieser Proteine des gewünschten Biosyntheseweges und damit durch einem verstärkten Stoffwechselfluß zu dem gewünschten biosynthetischen Produkt führt.

- 5 Je nach gewünschtem biosynthetischen Produkt muss die Transkriptionsrate bzw. Expressionsrate verschiedener Gene erhöht bzw. reduziert werden. In der Regel ist es vorteilhaft, die Transkriptionsrate bzw. Expressionsrate mehrere Gene zu verändern, d.h. die Transkriptionsrate bzw. Expressionsrate einer Kombination von Gene zu erhöhen und/oder die Transkriptionsrate bzw. Expressionsrate einer Kombination von
- 10 Gene zu reduzieren.

In den erfindungsgemäßigen genetisch veränderten Pflanzen ist mindestens eine erhöhte oder verursachte Expressionsrate eines Gens auf einen erfindungsgemäßigen Promotor zurückzuführen.

- 15 Weitere, zusätzliche veränderte, d.h. zusätzlich erhöhte oder zusätzlich reduzierte Expressionsraten von weiteren Genen in genetisch veränderten Pflanzen können, müssen aber nicht auf die erfindungsgemäßigen Promotoren zurück gehen.
- 20 Die Erfindung betrifft daher ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von erfindungsgemäßigen, genetisch veränderten Pflanzen der Gattung Tagetes.

Die Erfindung betrifft insbesondere ein Verfahren zur Herstellung von Carotinoiden

- 25 durch Kultivierung von erfindungsgemäßigen genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ϵ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein
- 30
- 35

FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Die Carotinoide sind vorzugsweise ausgewählt aus der Gruppe Phytoen, Phytofluen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-

5 Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Insbesondere betrifft die Erfindung weiterhin ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von erfindungsgemäß genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene

- 10 ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ϵ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase,
- 15 Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase,
- 20 Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.
- 25 Die Ketocarotinoide sind vorzugsweise ausgewählt aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Im erfindungsgemäß Verfahren zur Herstellung von biosynthetischen Produkten,

- 30 insbesondere Carotinoiden, vorzugsweise Ketocarotinoiden, wird vorzugsweise dem Kultivierungsschritt der genetisch veränderten Pflanzen ein Ernten der Pflanzen und ein Isolieren der biosynthetischen Produkte, insbesondere Carotinoide, vorzugsweise Ketocarotinoide aus den Pflanzen, vorzugsweise aus den Petalen der Pflanzen, ange- schlossen.

35

Die genetisch veränderten Pflanzen der Gattung Tagetes werden in an sich bekannter Weise auf Nährböden gezogen und entsprechend geerntet.

Die Isolierung von Ketocarotinoiden aus den geernteten Blütenblättern erfolgt bei-

- 40 spielsweise in an sich bekannter Weise, beispielsweise durch Trocknung und an-

schließender Extraktion und gegebenenfalls weiterer chemischer oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie. Die Isolierung von Ketocarotinoiden aus den Blütenblättern erfolgt beispielsweise bevorzugt durch organische Lösungsmittel wie Aceton, Hexan, Heptan, Ether oder tert.-Methylbutylether.

5 Weitere Isolierverfahren von Ketocarotinoiden, insbesondere aus Blütenblättern, sind beispielsweise in Egger und Kleinig (Phytochemistry (1967) 6, 437-440) und Egger (Phytochemistry (1965) 4, 609-618) beschrieben.

10 Ein besonders bevorzugtes Ketocarotinoid ist Astaxanthin.

15 Die Ketocarotinoide fallen im erfindungsgemäßen Verfahren in Blütenblättern in Form ihrer Mono- oder Diester mit Fettsäuren an. Einige nachgewiesene Fettsäuren sind z.B. Myristinsäure, Palmitinsäure, Stearinsäure, Ölsäure, Linolensäure, und Laurinsäure (Kamata und Simpson (1987) Comp. Biochem. Physiol. Vol. 86B(3), 587-591).

20 Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen oder Pflanzenteile, wie insbesondere Blütenblätter mit erhöhtem Gehalt an biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin, können auch beispielsweise direkt oder nach an sich bekannter Prozessierung als Nahrungsmittel oder Futtermittel oder als Futter- und Nahrungsergänzungsmittel verwendet werden.

25 Ferner können die genetisch veränderten Pflanzen zur Herstellung von biosynthetischen Produkt-, insbesondere Carotinoid-, insbesondere Ketocarotinoid-, insbesondere Astaxanthin-haltigen Extrakten und/oder zur Herstellung von Futter- und Nahrungsergänzungsmitteln sowie von Kosmetika und Pharmazeutika verwendet werden.

30 Die genetisch veränderten Pflanzen der Gattung Tagetes weisen im Vergleich zum Wildtyp einen erhöhten Gehalt an dem gewünschten biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin auf.

35 Unter einem erhöhten Gehalt wird in diesem Fall auch ein verursachter Gehalt an Ketocarotinoiden, bzw. Astaxanthin verstanden.

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

Allgemeine Experimentelle Bedingungen:
Sequenzanalyse rekombinanter DNA

- Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-
- 5 DNA-Sequenzierer der Firma Licor (Vertrieb durch MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Beispiel 1:
Amplifikation einer DNA, die die gesamte Primärsequenz der NOST-Ketolase aus
10 *Nostoc sp. PCC 7120* kodiert

Die DNA, die für die NOST-Ketolase aus *Nostoc sp. PCC 7120* kodiert, wurde mittels PCR aus *Nostoc sp. PCC 7120* (Stamm der "Pasteur Culture Collection of Cyanobacterium") amplifiziert.

15 Für die Präparation von genetischer DNA aus einer Suspensionskultur von *Nostoc sp. PCC 7120*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO₃, 0.04 g/l K₂PO₄·3H₂O, 0.075 g/l MgSO₄·H₂O, 0.036 g/l CaCl₂·2H₂O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l ED-

20 TA disodium magnesium, 0.04 g/l Na₂CO₃, 1ml trace metal mix A5+Co (2.86 g/l H₃BO₃, 1.81 g/l MnCl₂·4H₂O, 0.222 g/l ZnSO₄·7H₂O, 0.39 g/l NaMoO₄·2H₂O, 0.079 g/l CuSO₄·5H₂O, 0.0494 g/l Co(NO₃)₂·6H₂O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

25 Protokoll für DNA Isolation aus *Nostoc PCC7120*:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10minütige Zentrifugation bei 8 000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem
30 Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris HCl (pH 7.5) resuspendiert und in ein Eppendorf Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C

gelöst.

Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc PCC 7120*, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc sp. PCC 7120* unter Verwendung eines

- 5 sense-spezifischen Primers (NOSTF, SEQ ID No. 79) und eines antisense-spezifischen Primers (NOSTG SEQ ID No. 80) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- 10 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:

- 15 - 1 ul einer *Nostoc sp. PCC 7120* DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NOSTF (SEQ ID No. 79)
- 0.2 mM NOSTG (SEQ ID No. 80)
- 5 ul 10X PCR-Puffer (TAKARA)
- 0.25 ul R-Taq Polymerase (TAKARA)
- 20 - 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

- 25 35X94°C 1 Minute

55°C 1 Minuten

72°C 3 Minuten

1X72°C 10 Minuten

- 30 Die PCR-Amplifikation mit SEQ ID No. 79 und SEQ ID No. 80 resultierte in einem 805 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 81). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pGEM-T (Promega) kloniert und der Klon pNOSTF-G erhalten.

35

Sequenzierung des Klons pNOSTF-G mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 88,886-89,662 des Datenbankeintrages AP003592 identisch ist. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc sp. PCC 7120*.

Beispiel 2

Amplifikation einer DNA, die die gesamte Primärsequenz der NP196-Ketolase aus

5 *Nostoc punctiforme ATCC 29133* kodiert

Die DNA, die für die NP196-Ketolase aus *Nostoc punctiforme ATCC 29133* kodiert, wurde mittels PCR aus *Nostoc punctiforme ATCC 29133* (Stamm der "American Type Culture Collection") amplifiziert.

10

Für die Präparation von genomicscher DNA aus einer Suspensionskultur von *Nostoc punctiforme ATCC 29133*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO₃, 0.04 g/l K₂PO₄·3H₂O, 0.075 g/l MgSO₄·H₂O, 0.036 g/l CaCl₂·2H₂O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l EDTA disodium magnesium, 0.04 g/l Na₂CO₃, 1ml Trace Metal Mix "A5+Co" (2.86 g/l H₃BO₃, 1.81 g/l MnCl₂·4H₂O, 0.222 g/l ZnSO₄·7H₂O, 0.39 g/l Na-MoO₄·2H₂O, 0.079 g/l CuSO₄·5H₂O, 0.0494 g/l Co(NO₃)₂·6H₂O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

20

Protokoll für die DNA-Isolation aus *Nostoc punctiforme ATCC 29133*:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem

25 Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris-HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

30
35
35 Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme ATCC 29133*, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme ATCC 29133* unter Verwendung eines sense-spezifischen Primers (NP196-1, SEQ ID No. 82) und

eines antisense-spezifischen Primers (NP196-2 SEQ ID No. 83) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- 5 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:
- 1 ul einer *Nostoc punctiforme* ATCC 29133 DNA (hergestellt wie oben beschrieben)
 - 10 - 0.25 mM dNTPs
 - 0.2 mM NP196-1 (SEQ ID No. 82)
 - 0.2 mM NP196-2 (SEQ ID No. 83)
 - 5 ul 10X PCR-Puffer (TAKARA)
 - 0.25 ul R Taq Polymerase (TAKARA)
 - 15 - 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten
 - 20 35X 94°C 1 Minute
55°C 1 Minuten
72°C 3 Minuten
 - 1X72°C 10 Minuten
- 25 Die PCR-Amplifikation mit SEQ ID No. 82 und SEQ ID No. 83 resultierte in einem 792 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP196, SEQ ID No. 84). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP196 erhalten.
- 30 Sequenzierung des Klons pNP196 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 140.571-139.810 des Datenbank-eintrages NZ_AABC01000196 identisch ist (inverse orientiert zum veröffentlichten Datenbankeintrag) mit der Ausnahme, daß G in Position 140.571 durch A ersetzt wurde,
- 35 um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme* ATCC 29133.
- Dieser Klon pNP196 wurde daher für die Klonierung in den Expressionsvektor
- 40 pJIT117(Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

pJIT117 wurde modifiziert, indem der 35S-Terminator durch den OCS-Terminator (Octopine Synthase) des Ti-Plasmides pTi15955 von Agrobacterium tumefaciens (Datenbankeintrag X00493 von Position 12,541-12,350, Gielen et al. (1984) EMBO J. 3 835-846) ersetzt wurde.

Das DNA-Fragment, das die OCS-Terminatorregion beinhaltet, wurde mittels PCR unter Verwendung des Plasmides pHELLSGATE (Datenbankeintrag AJ311874, Wesley et al. (2001) Plant J. 27 581-590, nach Standardmethoden aus *E.coli* isoliert) sowie der Primer OCS-1 (SEQ ID No. 85) und OCS-2 (SEQ ID No. 86) hergestellt.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die die Octopin Synthasē (OCS) Terminatorregion (SEQ ID No. 87) beinhaltet, erfolgte in einem 50 μ l Reaktionsansatz, in dem enthalten waren:

- 100 ng pHELLSGATE plasmid DNA
- 0.25 mM dNTPs
- 0.2 mM OCS-1 (SEQ ID No. 85)
- 0.2 mM OCS-2 (SEQ ID No. 86)
- 5 μ l 10X PCR-Puffer (Stratagene)
- 0.25 μ l Pfu Polymerase (Stratagene)
- 28.8 μ l Aq. Dest.

25

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten
35X 94°C 1 Minute
30 50°C 1 Minute
72°C 1 Minute
1X72°C 10 Minuten

Das 210 bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pOCS erhalten.

Sequenzierung des Klons pOCS bestätigte eine Sequenz, die mit einem Sequenzabschnitt auf dem Ti-Plasmid pTi15955 von Agrobacterium tumefaciens (Datenbankeintrag X00493) von Position 12.541 bis 12.350 übereinstimmt.

40

Die Klonierung erfolgte durch Isolierung des 210 bp Sall-Xhol Fragmentes aus pOCS und Ligierung in den Sall-Xhol geschnittenen Vektor pJIT117.

Dieser Klon heisst pJO und wurde daher für die Klonierung in den Expressionsvektor pJONP196 verwendet.

5

Die Klonierung erfolgte durch Isolierung des 782 Bp SphI-Fragmentes aus pNP196 und Ligierung in den SphI geschnittenen Vektor pJO. Der Klon, der die NP196-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP196.

10

Beispiel 3:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Tagetes erecta*

15

Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus Petunia hybrida (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).

20

Das DNA Fragment, das die EPSPS Promoterregion (SEQ ID No. 88) aus Petunia hybrida beinhaltet, wurde mittels PCR unter Verwendung genomicscher DNA (nach Standardmethoden aus Petunia hybrida isoliert) sowie der Primer EPSPS-1 (SEQ ID No. 89) und EPSPS-2 (SEQ ID No. 90) hergestellt.

25

Die PCR-Bedingungen waren die folgenden:

30

Die PCR zur Amplifikation der DNA, die das EPSPS-Promoterfragment (Datenbankeintrag M37029: Nukleotidregion 7-1787) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 100 ng genomicscher DNA aus *A.thaliana*
- 0.25 mM dNTPs
- 0.2 mM EPSPS-1 (SEQ ID No. 89)
- 0.2 mM EPSPS-2 (SEQ ID No. 90)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

35X 94°C 1 Minute

5 50°C 1 Minute

72°C 2 Minute

1X72°C 10 Minuten

Das 1773 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-
10 Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pEPSPS erhalten.

Sequenzierung des Klons pEPSPS bestätigte eine Sequenz, die sich lediglich durch
zwei Deletion (Basen ctaagttcagga in Position 46-58 der Sequenz M37029; Basen
aaaaat in Position 1422-1429 der Sequenz M37029) und die Basenaustausche (T
15 statt G in Position 1447 der Sequenz M37029; A statt C in Position 1525 der Sequenz
M37029; A statt G in Position 1627 der Sequenz M37029) von der publizierten EPSPS-
Sequenz (Datenbankeintrag M37029: Nukleotidregion 7-1787) unterscheidet. Die zwei
Deletionen und die zwei Basenaustausche an den Positionen 1447 und 1627 der Se-
quenz M37029 wurden in einem unabhängigen Amplifikationsexperiment reproduziert
20 und repräsentieren somit die tatsächliche Nukleotidsequenz in den verwendeten Petu-
nia hybrida Pflanzen.

Der Klon pEPSPS wurde daher für die Klonierung in den Expressionsvektor pJONP196
verwendet.

25 Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus
pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJONP196. Der Klon,
der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heißt
pJOESP:NP196. Diese Expressionskassette enthält das Fragment NP196 in der kor-
rekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.
30

Zur Herstellung des Expressionsvektors MSP107 wurde das 2.961 KB bp SacI-Xhol
Fragment aus pJOESP:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN3 ligiert
MSP 107 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment rbcS
35 TP FRAGMENT das rbcS Transitpeptid aus Erbse (194 bp), Fragment NP196 KETO
CDS (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment OCS
Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 5 Zur Herstellung des Expressionsvektors MSP108 wurde das 2.961 KB bp SacI-Xhol Fragment aus pJOESP:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 108 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase , Fragment OCS 10 *Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Beispiel 4:

Amplifikation einer DNA, die die gesamte Primärsequenz der NP195-Ketolase aus *Nostoc punctiforme* ATCC 29133 kodiert

- 15 Die DNA, die für die NP195-Ketolase aus *Nostoc punctiforme* ATCC 29133 kodiert, wurde mittels PCR aus *Nostoc punctiforme* ATCC 29133 (Stamm der "American Type Culture Collection") amplifiziert. Die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc punctiforme* ATCC 29133 wurde in Beispiel 19 beschrieben.

- 20 Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme* ATCC 29133, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme* ATCC 29133 unter Verwendung eines sense-spezifischen Primers (NP195-1, SEQ ID No. 91) und eines antisense-spezifischen Primers (NP195-2 SEQ ID No. 92) amplifiziert.

- 25 25 Die PCR-Bedingungen waren die folgenden:

- 30 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:

- 1 ul einer *Nostoc punctiforme* ATCC 29133 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NP195-1 (SEQ ID No. 91)
- 0.2 mM NP195-2 (SEQ ID No. 92)
- 5 ul 10X PCR-Puffer (TAKARA)
- 0.25 ul R Taq Polymerase (TAKARA)
- 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten
35X 94°C 1 Minute
5 55°C 1 Minuten
72°C 3 Minuten
1X72°C 10 Minuten

- Die PCR-Amplifikation mit SEQ ID No. 91 und SEQ ID No. 92 resultierte in einem 819 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP195, SEQ ID No. 93). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP195 erhalten.
- Sequenzierung des Klons pNP195 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 55,604-56,392 des Datenbank-eintrages NZ_AABC010001965 identisch ist, mit der Ausnahme, daß T in Position 55.604 durch A ersetzt wurde, um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme* ATCC 29133.

Dieser Klon pNP195 wurde daher für die Klonierung in den Expressionsvektor pJO (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 809 Bp SphI-Fragments aus pNP195 und Ligierung in den SphI geschnittenen Vektor pJO. Der Klon, der die NP195-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP195.

Beispiel 5:
Amplifikation einer DNA, die die gesamte Primärsequenz der NODK-Ketolase aus *Nodularia spumignea* NSOR10 kodiert.

Die DNA, die für die Ketolase aus *Nodularia spumignea* NSOR10 kodiert, wurde mittels PCR aus *Nodularia spumignea* NSOR10 amplifiziert.

Für die Präparation von genomicscher DNA aus einer Suspensionskultur von *Nodularia spumignea* NSOR10, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO₃, 0.04 g/l K₂PO₄·3H₂O, 0.075 g/l MgSO₄·H₂O, 0.036 g/l CaCl₂·2H₂O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l

EDTA disodium magnesium, 0.04 g/l Na₂CO₃, 1ml Trace Metal Mix "A5+Co" (2.86 g/l H₃BO₃, 1.81 g/l MnCl₂·4H₂O, 0.222 g/l ZnSO₄·7H₂O, 0.39 g/l NaMoO₄·2H₂O, 0.079 g/l CuSO₄·5H₂O, 0.0494 g/l Co(NO₃)₂·6H₂O) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

5

Protokoll für die DNA-Isolation aus *Nodularia spumignea* NSOR10:

- Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem
10 Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris_HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von
100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert.
15 Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.
20

- Die Nukleinsäure, kodierend eine Ketolase aus *Nodularia spumignea* NSOR10, wurde mittels "polymerase chain reaction" (PCR) aus *Nodularia spumignea* NSOR10 unter Verwendung eines sense-spezifischen Primers (NODK-1, SEQ ID No. 94) und eines
25 antisense-spezifischen Primers (NODK-2 SEQ ID No. 95) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der
30 gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:

- 1 ul einer *Nodularia spumignea* NSOR10 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NODK-1 (SEQ ID No. 94)
- 0.2 mM NODK-2 (SEQ ID No. 95)
- 5 ul 10X PCR-Puffer (TAKARA)
- 0.25 ul R Taq Polymerase (TAKARA)
- 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten
35X 94°C 1 Minute
5 55°C 1 Minuten
72°C 3 Minuten
1X72°C 10 Minuten

- Die PCR-Amplifikation mit SEQ ID No. 94 und SEQ ID No. 95 resultierte in einem 720 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NODK, SEQ ID No. 96). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNODK erhalten.
- Sequenzierung des Klons pNODK mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 2130-2819 des Datenbank-eintrages AY210783 identisch ist (inverse orientiert zum veröffentlichten Datenbankeintrag). Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nodularia spumignea NSOR10*.

- Dieser Klon pNODK wurde daher für die Klonierung in den Expressionsvektor pJO (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 710 Bp SphI-Fragmentes aus pNODK und Ligierung in den SphI geschnittenen Vektor pJO.
- Der Klon, der die NODK-Ketolase von *Nodularia spumignea* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONODK.
- Beispiel 6:
Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NODK-Ketolase aus *Nodularia spumignea NSOR10* in *Lycopersicon esculentum* und *Tagetes erecta*.
- Die Expression der NODK-Ketolase aus *Nodularia spumignea NSOR10* in *L. esculentum* und *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus Petunia hybrida (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).

Der Klon pEPSPS (in Beispiel 8 beschrieben) wurde daher für die Klonierung in den Expressionsvektor pJONODK (in Beispiel 12 beschrieben) verwendet.

- Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJ0NODK. Der Klon, der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heisst pJOESP:NODK. Diese Expressionskassette enthält das Fragment NODK in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.
- 10 Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus *Nodularia spumignea NSOR10* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).
- 15 Zur Herstellung des Expressionsvektors MSP115 wurde das 2.889 KB bp SacI-Xhol Fragment aus pJOESP:NODK mit dem SacI-Xhol geschnittenen Vektor pSUN3 ligiert. MSP 115 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea NSOR10* NODK-Ketolase,
- 20 Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus *Nodularia spumignea NSOR10* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).
- Zur Herstellung des Expressionsvektors MSP116 wurde das 2.889 KB bp SacI-Xhol Fragment aus pJOESP:NODK mit dem SacI-Xhol geschnittenen Vektor pSUN5 ligiert.
- 30 MSP 116 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea NSOR10* NODK-Ketolase, Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 35 Beispiel 6A:
Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme ATCC 29133* in *Lycopersicon esculentum* und *Tagetes erecta*.

Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters PDS (Phytoendesaturase) aus *Lycopersicon esculentum* (Datenbankeintrag U46919).

5

Das DNA Fragment, das die PDS Promoterregion (SEQ ID No. 100) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomicscher DNA (nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer PDS-1 (SEQ ID No. 98) und PDS-2 (SEQ ID No. 99) hergestellt.

10

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das PDS-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

15

- 100 ng genomicscher DNA aus *Lycopersicon esculentum*
- 0.25 mM dNTPs
- 0.2 mM PDS-1 (SEQ ID No. 98)
- 0.2 mM PDS-2 (SEQ ID No. 99)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

20

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

25

- | | | | |
|-----|------|-----------|------------|
| 1X | 94°C | 2 Minuten | |
| 35X | 94°C | 1 Minute | |
| | 50°C | 1 Minute | |
| | 72°C | 2 Minuten | |
| 30 | 1X | 72°C | 10 Minuten |

Das 2096 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pPDS erhalten.

35

Der Klon pPDS wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 3 beschrieben) verwendet.

40

Die Klonierung erfolgte durch Isolierung des 2094 Bp Ecl136II-SmaI Fragmentes aus pPDS und Ligierung in den Ecl136II-HindIII geschnittenen Vektor pJOEPS:NP196. Die HindIII-Schnittstelle des Vektors wurde zuvor durch Behandlung mit dem Klenow-

Enzym in eine „blunt-end“-Schnittstelle überführt. Der Klon, der den Promoter PDS anstelle des ursprünglichen Promoters EPSPS enthält, heißt pJOPDS:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

5

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

- 10 Zur Herstellung des Expressionsvektors MSP117 wurde das 3.3 KB Ecl136II-Xhol Fragment aus pJOPDS:NP196 mit dem Ecl136II-Xhol geschnittenen Vektor pSUN3 ligiert. MSP 117 beinhaltet Fragment PDS den PDS Promoter, Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 15

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

20

Zur Herstellung des Expressionsvektors MSP118 wurde das 3.3 KB bp Ecl136II-Xhol Fragment aus pJOPDS:NP196 mit dem Ecl136II-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 118 beinhaltet Fragment PDS den PDS Promoter, Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase , Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Beispiel 6B:

- Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Lycopersicon esculentum* und *Tagetes erecta*.

Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715).

- 35 Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters B-GENE (chromoplastenspezifische lycopene B-cyclase) aus *Lycopersicon esculentum* (Datenbankeintrag AAZ51517).

Das DNA Fragment, das die B-GENE Promoterregion (SEQ ID No. 103) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomicscher DNA

40

(nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer BGEN-1 (SEQ ID No. 101) und BGEN-2 (SEQ ID No. 102) hergestellt.

Die PCR-Bedingungen waren die folgenden:

5

Die PCR zur Amplifikation der DNA, die das B-GENE-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 100 ng genomicscher DNA aus *Lycopersicon esculentum*

10

- 0.25 mM dNTPs

- 0.2 mM BGEN-1 (SEQ ID No. 101)

- 0.2 mM BGEN-2 (SEQ ID No. 102)

- 5 µl 10X PCR-Puffer (Stratagene)

- 0.25 µl Pfu Polymerase (Stratagene)

15

- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X 94°C 2 Minuten

20

35X 94°C 1 Minute

50°C 1 Minute

72°C 2 Minute

1X 72°C 10 Minuten

25

Das 1222 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pB-GENE erhalten.

Der Klon pB-GENE wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 3 beschrieben) verwendet.

30

Die Klonierung erfolgte durch Isolierung des 1222 Bp SacI-HindIII Fragmentes aus pB-GENE und Ligierung in den SacI-HindIII geschnittenen Vektor pJOEPS:NP196. Der Klon, der den Promoter B-GENE anstelle des ursprünglichen Promoters EPSPS enthält, heißt pJOBGEN:NP196. Diese Expressionskassette enthält das Fragment

35

NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

40

Zur Herstellung des Expressionsvektors MSP119 wurde das 2.4 KB SacI-Xhol Fragment aus pJOBGEN:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN3 ligiert. MSP 119 beinhaltet Fragment B-GENE den B-GENE Promoter, Fragment *rbcS TP*

- 5 *FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

- Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 10 Zur Herstellung des Expressionsvektors MSP120 wurde das 2.4 KB bp SacI-Xhol Fragment aus pJOBGEN:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 120 beinhaltet Fragment B-GENE den B-GENE Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase , Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

- 20 Beispiel 6C:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Lycopersicon esculentum* und *Tagetes erecta*.

- 25 Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters CHRC (chromoplast-specific carotenoid-associated protein) aus *Cucumis sativa* (Datenbank-eintrag AF099501).

30

Das DNA Fragment, das die CHRC Promoterregion (SEQ ID No. 106) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer CHRC-1 (SEQ ID No. 104) und CHRC-2 (SEQ ID No. 105) hergestellt.

35

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das CHRC-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

40

- 100 ng genomischer DNA aus *Lycopersicon esculentum*
- 0.25 mM dNTPs
- 0.2 mM CHRC-1 (SEQ ID No. 101)
- 0.2 mM CHRC-2 (SEQ ID No. 102)
- 5 - 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 10
- 1X 94°C 2 Minuten
 - 35X 94°C 1 Minute
 - 50°C 1 Minute
 - 72°C 2 Minute
 - 15 1X 72°C 10 Minuten

Das 1222 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pCHRC erhalten.

- 20 Der Klon pB-GENE wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 2 beschrieben) verwendet.

Die Klonierung erfolgte durch Isolierung des 1540 Bp SacI-HindIII Fragmentes aus pCHRC und Ligierung in den SacI-HindIII geschnittenen Vektor pJOEPS:NP196. Der 25 Klon, der den Promoter CHRC anstelle des ursprünglichen Promoters EPSPS enthält, heisst pJOCHRC:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

- 30 Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

Zur Herstellung des Expressionsvektors MSP121 wurde das 2.6 KB SacI-Xhol Fragment aus pJOCHRC:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN3 ligiert. 35 MSP 121 beinhaltet Fragment CHRC den CHRC Promoter, Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 5 Zur Herstellung des Expressionsvektors MSP122 wurde das 2.6 KB bp SacI-Xhol Fragment aus pJOCHRC:NP196 mit dem SacI-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 122 beinhaltet Fragment CHRC den CHRC Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase , Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 10

Beispiel 7:

Herstellung transgener Tagetes Pflanzen

- 15 Tagetessamen werden sterilisiert und auf Keimungsmedium (MS-Medium; Murashige and Skoog, Physiol. Plant. 15(1962), 473-497) pH 5,8, 2% Saccharose) aufgelegt. Die Keimung erfolgt in einem Temperatur/Licht/Zeitintervall von 18-28°C/20-200 µE/3 - 16 Wochen, bevorzugt jedoch bei 21°C, 20-70 µE, für 4-8 Wochen.
- 20 Alle Blätter der sich bis dahin entwickelten *in vitro* Pflanzen werden geerntet und quer zur Mittelrippe geschnitten. Die dadurch entstehenden Blattexplantate mit einer Größe von 10 - 60 mm² werden im Verlaufe der Präparation in flüssigem MS - Medium bei Raumtemperatur für maximal 2 h aufbewahrt.
- 25 Ein beliebiger Agrobakterium tumefaciens Stamm, bevorzugt aber ein supervirulenter Stamm, wie z.B. EHA105 mit einem entsprechenden Binärplasmid, das ein Selektionsmarkergen (bevorzugt *bar* oder *pat*) sowie ein oder mehrere Trait- oder Reporter-gene tragen kann wird (pS5FNR:NOST, pS5AP3:NOST pS5FNR:NP196, pS5EPS:NP196, pS5FNR:NP195, pS5EPS:NP195, pS5FNR:NODK und pS5EPS:NODK), über Nacht angezogen und für die Co-Kultivierung mit dem Blattmaterial verwendet. Die Anzucht des Bakterienstammes kann wie folgt erfolgen: Eine Einzelkolonie des entsprechenden Stammes wird in YEB (0,1 % Hefeextrakt, 0,5 % Rindfleischextrakt, 0,5 % Pepton, 0,5 % Saccharose, 0,5 % Magnesiumsulfat x 7 H₂O) mit 25 mg/l Kanamycin angeimpft und bei 28°C für 16 bis 20 h angezogen. Anschließend wird die Bakteriensuspension durch Zentrifugation bei 6000 g für 10 min geerntet und derart in flüssigem MS Medium resuspendiert, daß eine OD₆₀₀ von ca. 0,1 bis 0,8 entstand. Diese Suspension wird für die C-Kultivierung mit dem Blattmaterial verwendet.
- 35
- 40 Unmittelbar vor der Co-Kultivierung wird das MS-Medium, in dem die Blätter aufbewahrt worden sind, durch die Bakteriensuspension ersetzt. Die Inkubation der Blät-

chen in der Agrobakteriensuspension erfolgte für 30 min unter leichtem Schütteln bei Raumtemperatur. Anschließend werden die infizierten Explantate auf ein mit Agar (z.B. 0,8 % Plant Agar (Duchefa, NL) verfestigtes MS-Medium mit Wachstumsregulatoren, wie beispielsweise 3 mg/l Benzylaminopurin (BAP) sowie 1 mg/l Indolylessigsäure (IAA) aufgelegt. Die Orientierung der Blätter auf dem Medium ist bedeutungslos. Die Kultivierung der Explantate findet für 1 bis 8 Tage, bevorzugt aber für 6 Tage statt, dabei können folgende Bedingungen angewendet werden: Lichtintensität: 30 – 80 $\mu\text{Mol}/\text{m}^2 \times \text{sec}$, Temperatur: 22 – 24°C, hell/dunkel Wechsel von 16/8 Stunden. Anschließend werden die co-kultivierten Explantate auf frisches MS-Medium, bevorzugt mit den gleichen Wachstumsregulatoren übertragen, wobei dieses zweite Medium zusätzlich ein Antibiotikum zur Unterdrückung des Bakterienwachstums enthält. Timentin in einer Konzentration von 200 bis 500 mg/l ist für diesen Zweck sehr geeignet. Als zweite selektive Komponente wird eine für die Selektion des Transformationserfolges eingesetzt. Phosphinothricin in einer Konzentration von 1 bis 5 mg/l selektiert sehr effizient, aber auch andere selektive Komponenten gemäß des zu verwendenden Verfahrens sind denkbar.

Nach jeweils ein bis drei Wochen erfolgt der Transfer der Explantate auf frisches Medium bis sich Sproßknospen und kleine Sprosse entwickeln, die dann auf das gleiche Basalmedium einschließlich Timentin und PPT oder alternative Komponenten mit Wachstumsregulatoren, nämlich z.B. 0,5 mg/l Indolylbuttersäure (IBA) und 0,5 mg/l Gibberellinsäure GA₃, zur Bewurzelung übertragen werden. Bewurzelte Sprosse können ins Gewächshaus überführt werden.

Zusätzlich zu der beschriebenen Methode sind folgende vorteilhafte Modifikationen möglich:

Bevor die Explantate mit den Bakterien infiziert werden, können sie für 1 bis 12 Tage, bevorzugt 3 - 4, auf das oben beschriebene Medium für die Co-Kultur vorinkubiert werden. Anschließend erfolgt die Infektion, Co-Kultur und selektive Regeneration wie oben beschrieben.

Der pH Wert für die Regeneration (normalerweise 5,8) kann auf pH 5,2 gesenkt werden. Dadurch wird die Kontrolle des Agrobakterienwachstums verbessert.

Die Zugabe von AgNO₃ (3 - 10 mg/l) zum Regenerationsmedium verbessert den Zustand der Kultur einschließlich der Regeneration selbst.

Komponenten, die die Phenolbildung reduzieren und dem Fachmann bekannt sind, wie z.B. Zitronensäure, Ascorbinsäure, PVP u.v.a.m., wirken sich positiv auf die Kultur aus.

- 5 Für das gesamte Verfahren kann auch flüssiges Kulturmedium Verwendung finden. Die Kultur kann auch auf handelsüblichen Trägern, die auf dem flüssigen Medium positioniert werden inkubiert werden.

Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

10

Mit pS5FNR:NOST wurde beispielsweise erhalten: MSP102-1, MSP102-2, MSP102-3,

Mit pS5AP3:NOST wurde beispielsweise erhalten: MSP104-1, MSP104-2, MSP104-3

15

Mit pS5FNR:NP196 wurde erhalten: MSP106-1, MSP106-2, MSP106-3

Mit pS5EPS:NP196 wurde erhalten: MSP108-1, MSP108-2, MSP108-3

Mit pS5FNR:NP195 wurde erhalten: MSP110-1, MSP110-2, MSP110-3

20

Mit pS5EPS:NP195 wurde erhalten: MSP112-1, MSP112-2, MSP112-3

Mit pS5FNR:NODK wurde erhalten: MSP114-1, MSP114-2, MSP114-3

25

Mit pS5EPS:NODK wurde erhalten: MSP116-1, MSP116-2, MSP116-3

Mit pS3PDS:NP196 wurde erhalten: MSP117-1, MSP117-2, MSP117-3

Mit pS5PDS:NP196 wurde erhalten: MSP118-1, MSP118-2, MSP118-3

30

Mit pS3CHRC:NP196 wurde erhalten: MSP119-1, MSP119-2, MSP119-3

Mit pS5CHRC:NP196 wurde erhalten: MSP120-1, MSP120-2, MSP120-3

Mit pS3BGEN:NP196 wurde erhalten: MSP121-1, MSP121-2, MSP121-3

Mit pS5BGEN:NP196 wurde erhalten: MSP122-1, MSP122-2, MSP122-3

35

Beispiel 8:

Enzymatische Lipase-katalysierte Hydrolyse von Carotinoideestern aus Pflanzenmaterial und Identifizierung der Carotinoide

40

Allgemeine Arbeitsvorschrift

- a) Gemörsertes Pflanzenmaterial (z.B. Petalenmaterial) (30-100 mg Frischgewicht) wird mit 100% Aceton (dreimal 500 μ l; jeweils etwa 15 Minuten schütteln) extrahiert.
- 5 Das Lösungsmittel wird evapiert. Carotinoide werden anschließend in 495 μ l Aceton aufgenommen, 4,95 ml Kaliumphosphatpuffer (100 mM, pH7.4) zugegeben und gut gemischt. Danach erfolgt die Zugabe von ca. 17 mg Bile-Salze (Sigma) und 149 μ l einer NaCl/CaCl₂-Lösung (3M NaCl und 75 mM CaCl₂). Die Suspension wird für 30 Minuten bei 37°C inkubiert. Für die enzymatische Hydrolyse der Carotinoidester wird 595 μ l einer Lipaselösung (50 mg/ml Lipase Typ7 von *Candida rugosa* (Sigma)) zugegeben und unter Schütteln bei 37C inkubiert. Nach etwa 21 Stunden erfolgte nochmals eine Zugabe von 595 μ l Lipase mit erneuter Inkubation von mindestens 5 Stunden bei 37°C. Anschließend werden etwa ca. 700 mg Na₂SO₄ in der Lösung gelöst. Nach Zugabe von 1800 μ l Petrolether werden die Carotinoide durch kräftig Mischen in die organische
- 10 Phase extrahiert. Dieses Ausschütteln wird solange wiederholt, bis die organische Phase farblos bleibt. Die Petroletherfraktionen werden vereinigt und der Petrolether evapiert. Freie Carotinoide werden in 100-120 μ l Aceton aufgenommen. Mittels HPLC und C30-reverse phase-Säule können freie Carotinoide aufgrund von Retentionszeit und UV-VIS-Spektren identifiziert werden.
- 15
- 20 Die verwendeten Bile-Salze oder Gallensäuresalze sind 1:1 Mischungen von Cholat und Desoxycholat.
- b) Arbeitsvorschrift für Aufarbeitung, wenn nur geringe Mengen an Carotinoidestern im
- 25 Pflanzenmaterial vorhanden sind
- Alternativ kann die Hydrolyse der Carotinoidester durch Lipase aus *Candida rugosa* nach Trennung mittels Dünnschichtchromatographie erreicht werden. Dazu werden 50-100mg Pflanzenmaterial dreimal mit etwa 750 μ l Aceton extrahiert. Der Lösungsmittel-
- 30 telextrakt wird im Vakuum einrotiert (erhöhte Temperaturen von 40-50°C sind tolerabel). Danach erfolgt Zugabe von 300 μ l Petrolether:Aceton (Verhältnis 5:1) und gute Durchmischung. Schwebstoffe werden durch Zentrifugation (1-2 Minuten) sedimentiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt. Das verbleibende Rest wird erneut mit 200 μ l Petrolether:Aceton (Verhältnis 5:1) extrahiert und Schwebstoffe
- 35 werden durch Zentrifugation entfernt. Die beiden Extrakte werden zusammengeführt (Volumen 500 μ l) und die Lösungsmittel evapiert. Der Rückstand wird in 30 μ l Petrolether:Aceton (Verhältnis 5:1) resuspendiert und auf eine Dünnschichtplatte (Silica-Gel 60, Merck) aufgetragen. Falls mehr als eine Auftragung für präparativ-analytische Zwecke erforderlich ist, sollten mehrere Aliquots mit jeweils 50-100 mg Frischgewicht in der

beschriebenen Weise für die dünnenschichtchromatographische Trennung aufbereitet werden.

- Die Dünnschichtplatte wird in Petrolether:Aceton (Verhältnis 5:1) entwickelt. Carotinoidbanden können visuell aufgrund ihrer Farbe identifiziert werden. Einzelne Carotinoidbanden werden ausgekratzt und können für präparativ-analytische Zwecke gepoolt werden. Mit Aceton werden die Carotinoide vom Silica-Material eluiert; das Lösungsmittel wird im Vakuum evaporiert. Zur Hydrolyse der Carotinoidester wird der Rückstand in 495µl Aceton gelöst, 17mg Bile-Salze (Sigma), 4,95ml 0.1M Kaliumphosphatpuffer (pH 7,4) und 149µl (3M NaCl, 75mM CaCl₂) zugegeben. Nach guter Durchmischung wird 30min bei 37°C äquilibriert. Danach erfolgt die Zugabe von 595µl Lipase von Candida rugosa (Sigma, Stammlösung von 50mg/ml in 5mM CaCl₂). Über Nacht erfolgt die Inkubation mit Lipase unter Schütteln bei 37°C. Nach etwa 21 Stunden wird nochmals die gleiche Menge an Lipase zugegeben; für mindestens 5 Stunden wird 15 nochmals bei 37°C unter Schütteln inkubiert. Dann erfolgt die Zugabe von 700mg Na₂SO₄ (wasserfrei); mit 1800µl Petrolether wird für ca. 1 Minute ausgeschüttelt und die Mischung bei 3500 Umdrehungen/Minute für 5 Minuten zentrifugiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt und das Ausschütteln so lange wiederholt, bis die obere Phase farblos ist. Die vereinigte Petrolether-Phase wird im Vakuum eingeengt (Temperaturen von 40-50°C sind möglich). Der Rückstand wird in 120µl Aceton, eventuell mittels Ultraschall, gelöst. Die gelösten Carotinoide können mittels HPLC unter Verwendung einer C30-Säule getrennt und anhand von Referenzsubstanzen quantifiziert werden.

25 Beispiel 9:

HPLC-Analyse freier Carotinoide

Die Analyse der nach der Arbeitsvorschriften in Beispiel 15 erhaltenen Proben erfolgt unter folgenden Bedingungen:

30 Folgende HPLC-Bedingungen wurden eingestellt.

Trennsäule: Prontosil C30-Säule, 250 x 4,6 mm, (Bischoff, Leonberg, Germany)

Flussrate: 1.0 ml/min

Eluenten: Laufmittel A - 100% Methanol

35 Laufmittel B - 80% Methanol, 0.2% Ammoniumacetat

Laufmittel C - 100% t-Butyl-methylether

Detektion: 300-530 nm

Gradientenprofil:

| Zeit | Flussrate | % Laufmittel A | % Laufmittel B | % Laufmittel C |
|------|-----------|----------------|----------------|----------------|
|------|-----------|----------------|----------------|----------------|

| | | | | |
|-------|-----|------|-----|------|
| 1.00 | 1.0 | 95.0 | 5.0 | 0 |
| 12.00 | 1.0 | 95.0 | 5.0 | 0 |
| 12.10 | 1.0 | 80.0 | 5.0 | 15.0 |
| 22.00 | 1.0 | 76.0 | 5.0 | 19.0 |
| 22.10 | 1.0 | 66.5 | 5.0 | 28.5 |
| 38.00 | 1.0 | 15.0 | 5.0 | 80.0 |
| 45.00 | 1.0 | 95.0 | 5.0 | 0 |
| 46.0 | 1.0 | 95.0 | 5.0 | 0 |

Einige typische Retentionszeiten für erfindungsgemäß gebildete Carotinoide sind z.B.: Violaxanthin 11,7 min, Astaxanthin 17,7 min, Adonixanthin 19 min, Adonirubin 19,9 min, Zeaxanthin 21 min.

Patentansprüche

1. Verwendung eines Promotors, ausgewählt aus der Gruppe
- 5 A) EPSPS Promotor
 B) B-Gene Promotor
 C) PDS Promotor und
 D) CHRC Promotor
- 10 zur Expression von Genen in Pflanzen der Gattung Tagetes, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
- 15 2. Verwendung nach Anspruch 1, dadurch gekennzeichnet, dass die Expression spezifisch in Blüten erfolgt.
3. Verwendung nach Anspruch 2, dadurch gekennzeichnet, dass die Expression spezifisch in Petalen erfolgt.
- 20 4. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der EPSPS Promotor gemäß Anspruch 1
 - A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder
 - A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder 3 aufweist oder
 - A3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert oder
 - A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3)
- 25 enthält.
- 30 5. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der B-Gene Promotor gemäß Anspruch 1
 - B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder

- B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist oder
- 5 B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder
- B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3)
- enthält.
- 10 6. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der PDS Promotor gemäß Anspruch 1
- C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder
- 15 C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder
- C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.
- 20 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder
- C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3)
- enthält.
- 25 7. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der CHRC Promotor gemäß Anspruch 1
- D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
- 30 D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
- D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.
- 35 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
- D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3)
- enthält.

8. Genetisch veränderte Pflanze der Gattung Tagetes, wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7.
9. Genetisch veränderte Pflanze nach Anspruch 8, dadurch gekennzeichnet, dass die Regulation der Expression von Genen in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7 dadurch erreicht wird, dass man
 - 10 a) eine oder mehrere Promotoren gemäß einem der Ansprüche 1 bis 7 in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer endogener Gene unter der Kontrolle der eingebrachten Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
 - 15 b) ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
 - 20 c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft eine oder mehrere, zu exprimierende Gene, in die Pflanze einbringt.
- 25 10. Genetisch veränderte Pflanze der Gattung Tagetes, enthaltend einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft ein zu exprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
- 30 11. Genetisch veränderte Pflanze nach einem der Ansprüche 8 bis 10, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nuk-

- leinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Kohlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Co-faktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen, wobei die Gene gegebenenfalls weitere Regulationselemente enthalten können.
- 10 12. Genetisch veränderte Pflanze nach Anspruch 11, dadurch gekennzeichnet, dass man als zu exprimierende Gene Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden verwendet.
- 15 13. Genetisch veränderte Pflanze nach Anspruch 12, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ε -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.
- 30 14. Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von genetisch veränderten Pflanzen der Gattung Tagetes gemäß einem der Ansprüche 8 bis 13.

15. Verfahren zur Herstellung von Carotinoiden durch Kultivierung von genetisch veränderten Pflanzen gemäß einem der Ansprüche 8 bis 13, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β -Hydroxylase, Nukleinsäuren kodierend eine β -Cyclase, Nukleinsäuren kodierend eine ε -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- Δ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.
16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass man nach dem Kultivieren die genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den genetisch veränderten Pflanzen isoliert.
17. Verfahren nach Anspruch 16, dadurch gekennzeichnet, dass man die Blüten der genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den Petalen der genetisch veränderten Pflanzen isoliert.
18. Verfahren nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, dass die Carotinoide ausgewählt sind aus der Gruppe Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

SEQUENCE LISTING

<110> SunGene GmbH & Co. KGaA

<120> Promotoren zur Expression von Genen in Tagetes

<130> PF 55341

<160> 106

<170> PatentIn version 3.1

<210> 1

<211> 174

<212> DNA

<213> Petunia hybrida

<220>

<221> promoter

<222> (1)..(174)

<223> EPSPS Promotor

<400> 1
gagaacacacag ctggaaatttt ttacaaagggt agttggtgaa gctagtcagc gaatcccatt 60
accttccact ctacctaacc cccttcacca acaacaaatt tctgtaattt aaaaacttagc 120
caaaaaagaa ctctctttta caaagagcca aagactcaat ctttactttc aaga 174

<210> 2

<211> 1781

<212> DNA

<213> Petunia hybrida

<220>

<221> promoter

<222> (1) ..(1781)

<223> EPSPS Promotor

<400> 2
tagcgcataatc ttatgtggta caaatcttga ttagtcgggc taagtttcag gaaaaaaaaatg 60
atgtggccctt acaaatacggtt ggaggatggg agatttggct ctatcttagag ttatgtggtt 120
gttgaagcat ttgggttactc tctgctgtgg tagttggcat atccacattg ttccttcca 180
cttttatgac aattacgtga aagttatggg ttgtttgtc tattttgtc gaggccttc 240
ttttccttcc aggttgttga agatggtcca attcgattag aataatgttt tgagctttag 300
catattctct ctcgtttaca cgattatagt aataatgata taggatgaca gaagttgaca 360
cataaaatttt ttattcttc catttacttt aatccaaatc tcacccatccc taaaacttctt 420
taatatgtat tcaatagtct atccgagtaa attgtaaatt taacaaccat tgataatatt 480
gacacctact aacatatact agtaaagaga atattaacat ggcacatata atttgatgca 540
aaatgagtat gatgaaattt aaacccaaaaa tctcttgatt ttgacagtgt caccttgact 600
tgttaactaa taagtcatgt ttttagtggca gaaagacaaa ctcatccacc aactgtatag 660
caataaaaaaa tagaagaatc ttccctgaggc aaagttttgg aaaaattaag agtggctgag 720
attnaatttc aacaggaatt agttccactt aacttttagg ttacgataca gtgctaatta 780
aataacttaa ttgtattaga tatttcttgc acctaaaaaa tttaaaaaact gaaaaaaaggt 840
agcaatcaaa ataaacaaaaa ggacaaaata agtggaaaggt acagccacca accctggcgg 900
ctcactgttt gttggttaaa acgttagactt acacccatcca aaatctacaa ctaaaatgag 960
gcaataatac tttgccaaaa attaccaaga aaagaaaaag aaaggaatcc cttaatatta 1020
ctctcctcca tttcacaata aatatcctag tttgacttaa attagagttt aaaaatgaa 1080
agacgacttt taaaacttgt aatctaaaat aaatcatagt taaatgtgtg gctataaatac 1140
attgtattaa cggtaaagtg gtaagttaa aagtttaattg tttcaataa taaaattgta 1200
ctatcattct ttttggaaatg gactaataag aaaactatga catccattat ggagcggagg 1260

gagtatctcc ttttaacaat aaccttgtc cttcaattc aattatcagt atgcaaacat 1320
taaaaattat tattgatgtt aagtaccaca tcatccttaa tgatagaatc atcgtagaac 1380
gctttccag gcacacattc aaactagttt gacaaaaaaa tatgtaccac acatcgaata 1440
gtccagactt cttgtttga atagtcgact acattggata atggaacttc tcgaattaac 1500
ttcgaattag tcgagcccc aataatatat acgtcggtt gaaaactata aatgtttga 1560
caaaaatgtc aaattaatat atcaatctgc aacaacctt tcaccttgag aacacagctg 1620
gaattttta caaaggttgtt tggtgaagct agtcagcgaa tcccattacc ttccactcta 1680
cctaaccccc ttcacccaaca acaaattct gtaattaaa aactagccaa aaaagaactc 1740
tctttacaa agagccaaag actcaatctt tactttcaag a 1781

<210> 3

<211> 1760

<212> DNA

<213> Petunia hybrida

<220>

<221> promoter

<222> (1)..(1760)

<223> EPSPS Promotor

<400> 3

tagcgcaatc ttatgtggta caaatcttga ttatcggtt aaaaatgtt tggccctaca 60
aatggtttga ggatgggaga tttggctcta tcttaggtt tgtggttttt gaagcatgg 120
gttactctct gctgtggtag ttggcatatc cacattgtct cttccactt ttatgacaat 180
tacgtgaaag ttatgggtt ttttgtctat ttttgtcgag gcctttctt tccttccagg 240
ttgttgaaga tggtccaatt cgattagaat aatgttttga gcttttagcat attctcttc 300
gtttacacga ttatagtaat aatgatatac gatgacagaa gttgacacat aaattttta 360
ttctctccat ttacttaat ccaaattctca cctaccctaa acttctttaa tatgtattca 420
atagtctatc cgagtaaattt gtaaattttaa caaccattga taatattgac acctactaac 480

atatactagt aaagagaata ttaacatggc acatataatt tgatgcaaaa tgagtatgtat 540
gaaatttaaa cccaaaatct cttgattttg acagtgtcac cttgacttgt taactaataa 600
gtcatgtttt agtggcagaa agacaaaactc atccaccaac tgtatagcaa taaaaaatag 660
aagaatcttc ctgaggcaaa gttttggaaa aattaagagt ggctgagatt taatttcaac 720
aggaatttagt tccacttaac ttttaggtta cgatacagtg ctaattaaat aacttaattg 780
tattagatat ttcttgacc taaaaaattt aaaaactgaa aaaaggttagc aatcaaaaata 840
aacaaaaagga caaaaataagt gaaaggtaca gccaccaacc ctggcggtc actgtttgtt 900
ggtaaaaacg tagacttaca cctaccaaaa tctacaacta aaatgaggca ataatacttt 960
gcccaaaatt accaagaaaa gaaaaagaaa ggaatccctt aatattactc tcctccattt 1020
cacaataaaat atccttagttt gacttaaattt agagttaaa aaatgaaaga cgacttttaa 1080
aacttgcataat ctAAAataaa tcatagttaa atgtgtggct ataaatcatt gtattaacgg 1140
taaagtggta agtttaaaag ttaattgttt tcaaataataa aattgtacta tcattcttt 1200
tggaatggac taataagaaa actatgacat ccattatgga gcggagggag tatctccctt 1260
taacaataac ctttgcctt tcaattcaat tattcgtatg caaacattaa aaattattat 1320
tgatgttaag taccacatca tccttaatga tagaatcatc gttagaacgct tttccaggca 1380
cacattcaaa ctagtttagac cagtagccaca catcgaatat tccagacttc tttgtttgaa 1440
tagtcgacta cattggataa tggaacttct cgaattaact tcgaatttagt cgagccccaa 1500
ataatatata cgtcgggtgg aaaactataa aatgtttgac aaaaatgtca aattaatata 1560
tcaatctgca acaacctttt caccttgaga acacagctga aattttttac aaaggttagtt 1620
ggtgaagcta gtcagcgaat cccattacct tccactctac ctaacccct tcaccaacaa 1680
caaatttctg taatttaaaa actagccaaa aaagaactct cttttacaaa gagccaaaga 1740
ctcaatcttt actttcaaga 1760

<210> 4

<211> 1210

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(1210)

<223> B-Gene Promotor

<400> 4

| | | | | | | |
|--------------|-------------|--------------|-------------|------------|--------------|------|
| gaattctctg | aaaaggagca | ccatattgc | cgcactgtgg | ttcatatttc | caagtacatt | 60 |
| tagatgaact | atatcatcag | attgaaagggt | tattgtataa | tcaatccagt | ggattctcggt | 120 |
| tctggcacct | ttagaagtac | atgtgcggaa | aagaatgata | aggtttgtat | tgttgttgac | 180 |
| aaagcctgtt | gcctttctca | tttgtaaatg | ttctgaacga | ctcctaaatt | actcttaagg | 240 |
| tgtaaggct | tccgtgcctg | tttgtaaata | taatgctgtg | ccgtgactta | cctttgtac | 300 |
| catttgttca | aatgtatggc | ctgaacacca | gggttgtcaa | aaatgtctca | tgcccgttt | 360 |
| attggctctga | aaatggcgtg | atgccaaattt | ctgcccgtcc | acagtgagca | tttcgatcta | 420 |
| ctggaaattt | accaacttat | tttattcactt | gataactaaa | caaaatccta | ttaactttaa | 480 |
| tcatacattt | tatttataacc | aaaaaaatttta | tgcataactc | attaaattac | cttttttagc | 540 |
| agtcaaattt | taaatcagtt | tctaattttat | caaaatggct | ttttaggggt | cccatattcca | 600 |
| ctaataacc | tgccgtccat | gcactgacta | caaaacaat | acctcactat | gtttgttagt | 660 |
| gttggtaat | ataaaaacctt | ttcttttatg | agaaaagttca | ccgagaataa | ttttctattt | 720 |
| gtggcataat | agtatatagt | gcagattgac | aagaattttaa | ttttgcagtt | gggcacatga | 780 |
| acaattttcc | tcaaagttgt | agaaagtact | tttcattttc | ttgtcaccga | aaatttatttta | 840 |
| taattgaaat | aaaaaccgaa | tgagctgcaa | gattcaagtc | gaattttcaa | aagaattgac | 900 |
| caagaaaaaaaa | ttcaaaaata | tcccccaccc | cctaccaaac | acatcctaaa | gtgaggtata | 960 |
| gactgggact | gggattggga | aaagggtaaa | atgctttcac | tagcttagca | aagattccac | 1020 |
| tttggtagct | atctttcttt | ctcatttcct | tttttttttt | tctttttttt | gttatataag | 1080 |
| ccaaagttagg | tacccaaaag | catcaatatt | ttgtattgct | tggtgattcc | tctgttagtcc | 1140 |
| agtattttcat | tttctacaag | ttccacctcc | ctccataatt | aaccattatc | aatcttatac | 1200 |
| attctctata | | | | | | 1210 |

<210> 5

<211> 1599

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(1599)

<223> B-Gene Promotor

| | |
|---|------|
| <400> 5 | |
| atctcattgt atagcttgtc ttttgtttca gtcgtcttag gcttgggtta gttggtgttg | 60 |
| ctgtttcata cttcttatcaa ccttgtgtga gttcctttat aaaatatgac tgttggagga | 120 |
| agtaatttac cttagttcg actacatcaa gatttgcate attctcgatcc aagaaatctt | 180 |
| agtttgaagc cttttggtct ggtatatttg tcaatctgag cttcgcaact ttctcatgac | 240 |
| aggggtttgt tgacatgctt gatttgctc ttcccttact tgataattgc tgcttggc | 300 |
| ggaggcatca ctctacccatc ctgcagatca tgaattctct gaaaaggagc accatatttg | 360 |
| ccgcactgtg gttcatattt ccaattacat ttagatgaac tatcatca ggagtgaaag | 420 |
| gttattgtat aatcaatcca gtggattctc gttctggcac cttagaagt acatgtgcgg | 480 |
| aaaagaatga taaggttgtt attgttggta acaaggcctg ttgccttct catttgtaaa | 540 |
| tgttctgaac gactcctaaa ttactctaa agtctaagggt ctccgtgcc tgttgtata | 600 |
| tataatgctg tgccgtgact tacctttgtt accattttgtt caaatgtatg gcctggacac | 660 |
| tagggttgtc aaaaatgtct catgacttca cccttcttc ttgtcttggt gcccgttttta | 720 |
| ttggctctgag aacggcgtga tgccaaattc tgccgctcca cagtggcat ttgcatctac | 780 |
| tggaaattga ccaacttatt ttatcacttg ataactagag tctgggttca aacaaaatcc | 840 |
| aataacttca atcatacatt gtatttatat tgaaaaattt atgcacaact cagtaaatta | 900 |
| cctttttttt cagtcaaaaa ttcttagatca gtttctaatt aatcaaaaatg gcctttatag | 960 |
| ggccccagtt ccattaatat acctgcccgtc catgcactga ttacaagaca aatacctcac | 1020 |
| tatgtttgtt agtgcgttggt aatataaaac cttttctttt atgagaaagt tcaccgaaaa | 1080 |
| taattttcttta tttgtggcat aactagtatc gaagtatata gtgcagattt acaagaattt | 1140 |

aatttgcag ttggcacat gaacaattt cctcaaagg ttagaaaata ttttcattt 1200
tcttgcacc gaaaattatt tataattgaa attgaaaccc aatgagctgc aagactcgag 1260
tcgaatttca aaaaaattga ccaactaaat atgaaaaaat ccgaatatat cccccacccc 1320
ctaccaaaca catcctaaag tgaggtatag actggactg ggattggaa aaggtaaaa 1380
tgcttcact agcttagcaa agatccact ttgttagcta tcttcttc tcatttcctt 1440
ttttctttt ctttttttt ttatataagc caaagtaggt accaaaaagc atcaatattt 1500
tgtattgctt ggtgattcct ctttactcca gtatttcatt ttctacaagt tccacccccc 1560
tccataatta accattatca atcttataca ttttctata 1599

<210> 6

<211> 1204

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(1204)

<223> B-Gene Promotor

<400> 6
tctgaaaagg agcaccatat ttgccgcact gtggttcata tttccaagta cattttagatg 60
aactatataca tcagattgaa aggttattgt ataatcaatc cagtggattc tcgttctggc 120
accttttagaa gtacatgtgc ggaaaagaat gataagggtt gtattgttgt tgacaaagcc 180
tgttgccttt ctcatttgta aatgttctga acgactccctt aattactctt aagggtgttaag 240
gtcttccgtg cctgtttgtta aatataatgc tgtgccgtga cttacccctt gtaccatgg 300
ttcaaatgtta tggcctggac accagggttg tcaaaaaatgt ctcatgcccc ttttatttgt 360
ctgagaatgg cgtgatgccaa aattctgccc ctccacagtgc agcatttcga tctactggaa 420
attgaccaac ttattttatc acttgataac taaacaaaat cctattaact ttaatcatac 480
attgtattta taccgaaaaaa gttatgcata actcagtaaa ttaccccttt tagcagtcaa 540

| | |
|--|------|
| attctagatc agtttcta at ttatcaaaat ggctttata gggcccagt tccactaata | 600 |
| tacctgccgt ccatgcactg actacaagac aaatacctca ctatgttgt tagtgcttgg | 660 |
| taatataaaa cctttcttt tatgagaaaag ttcaccgaga ataattttct atttgtggca | 720 |
| taactagtat atagtgcaga ttgacaagaa ttaattttc cagttggca catgaacaat | 780 |
| tttcctcaaa gttgttagaaa gtactttca ttttcttgc accgaaaatt atttataatt | 840 |
| gaaataaaaa ccgaatgagc tgcaagattc aagtcgaatt tcaaaaagaat tgaccaagaa | 900 |
| aaaattcaaa aatatcccccc accccctacc aaacacatcc taaagtgagg tatagactgg | 960 |
| gactgggatt gggaaaaggg taaaatgctt tcactagctt agcaaagatt ccactttgtt | 1020 |
| agctatcttt ctttctcatt tcctttttc tttttctttt ttttgttata taagccaaag | 1080 |
| taggtaccca aaagcatcaa tattttgtat tgcttggta ttccctctgta gtccagtatt | 1140 |
| tcattttcta caagttccac ctccctccat aattaaccat tatcaatctt atacattctc | 1200 |
| tata | 1204 |

<210> 7

<211> 2078

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1)..(2078)

<223> PDS Promotor

| | |
|--|-----|
| <400> 7 | |
| tttgcagta ttacaacagc ttatatgttg agcaggtaaa agcttcaatg ccctattttt | 60 |
| tctacagtta tcaatgttgc tcgtctaata tctgggttgc ttctcgaaat gtcaattggc | 120 |
| ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgc当地 | 180 |
| atttatataat ttcatagtg tcaatgttgc ttcttttagct ggttcctcat agtaaagttg | 240 |
| tctaataatga aatgaaaatg ttctgtgtgt tgtactaata cctttcatg gttgtctata | 300 |

| | |
|---|------|
| gaacgtcgat gaagagccaa acagaaaacta ttttgggctg cgatttctga taccattgta | 360 |
| tctgaatgct gggtgggagc tcatacagaag ctttacaatg ggtcacatata atggagccgg | 420 |
| tatgaggaat gctggaaatc agttgcgttt cgcgtgctag gacttttcct tcctggatt | 480 |
| tctgccaca gcccagttga ttacgtgaac tccgtcagac ttggaaagga gagaagtacc | 540 |
| caaatagttgt cttttagaa atactttgt cacaaaatag cggggtttac agctacagaa | 600 |
| gatcatgcag aaggcgtcca gtttagttt tgaaggttgt ttggagttta tttatctaaa | 660 |
| gtaaaactaa atcagcttt tgtttatgag ttcaagtgaac tataatgtca aataagactt | 720 |
| ccctttgttag atatgtgttt ttttgttgt tgagcacttt gtgtgcattt gataaacc | 780 |
| caacgtgtaa tagctaccat acaagagaag taactcgac tgtccatgtc ttatgtggct | 840 |
| cgactcagaa agcattcagg gggattgata accaccctcc aaaccaactg aaccatttg | 900 |
| aataaccacc cttcaaatca accgagtcct cgtgaaggac aaatatgtgg ttttatatac | 960 |
| attaaatttt gtttttacat gtttcctctt acttcttttag ttttcttgac catatctgc | 1020 |
| gtttttccct tctgttaattt acactttct tcaaaccatc cagcaatgtg gaagcttgac | 1080 |
| gattttcattt cagagttagaa attgaaaaga atcaactaaa aaggatagtc ttccgatttgc | 1140 |
| atttccggct taaaaataaa ctaataagaa tgagagagcg aataatagaa tattttgaaa | 1200 |
| ttttaaagat attcaactat gttaaattgc gttataaatt tcttaaatta gtagcaccta | 1260 |
| atagtttagt tctcaaaagt caaaaactact acataatgtg ctcattttc acattaaaat | 1320 |
| gcctacatga tgtaaaagta aaactcgtag cattctacgt gtttactca actcaaacat | 1380 |
| cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt tcttttttt | 1440 |
| ttttaaaaaaaa atatttttt ttatataat ccaaattggc tccaaatttat cataaattag | 1500 |
| gtagaaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg ctatggtgg | 1560 |
| acgtgtcaaa ttttggattt tagccaaaca tgagatttga tttaaaggaa attggccaaa | 1620 |
| tcaccgaaag caggcatctt catcataaat tagttgttt atttatacag aattatacgc | 1680 |
| ttttactagt tatagcattc ggtatcttt tctggtaac tgccaaacca ccacaaattt | 1740 |
| caagtttcca tttaactctt caacttcaac ccaaccaaattt ttatgtgcatt aattgtgcag | 1800 |
| aaccactccc tatatcttctt aggtgcatttcc attcgatccg aggtaaagaaa agatttttgt | 1860 |
| ttctttgaat gctttatgcc actcgatcaa cttctgaggt ttgtggatct tttaggcgac | 1920 |
| ttttttttt tttgtatgta aaatttgttt cataaaatgct tctcaacata aatcttgaca | 1980 |

aagagaagga attttaccaa gtat taggt tcagaaatgg ataatttct tactgtgaaa 2040
tatcctttag gcaggttta ctgttatttt tcagtaaa 2078

<210> 8

<211> 1342

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)...(1342)

<223> PDS Promotor

<400> 8
gttttttttg ttgtttagca ctttgtgtgc attggataaaa cccccaacgt gtaatagcta 60
ccataacaaga gaagtaactc gcactgtcca tgtcttatgt ggctcgactc agaaaggcatt 120
cagggggatt gataaccacc ctccaaacca actgaaccat tgtgaataac cacccttcaa 180
atcaaccgag tcctcgtaaa ggacaaatat gtggttttat atacattaaa ttttggaaaa 240
acatgcttcc tcttacttct ttagtttct tgaccatatc ttgcgtttt cccttctgta 300
attgacaccc ttcttcaaacc catccagcaa tgtgaaagct tgacgatttt ctttcagagt 360
agaaattgaa aagaatcaac taaaaaggat agtccttcga tttgatttcc ggcttaaaaa 420
taaactaata agaatgagag agcgaataat agaatatttt gaaattttaa agatattcaa 480
ctatgttaaa ttgcgttata aatttcttaa attagtagca cctaatacgat tagttctcaa 540
aagtcaaaac tactacataa tgtgctcatt tttcacatata aaatgcctac atgatgtaaa 600
agtaaaaactc gtagcattct acgtgtttta ctcaactcaa acatcctgtt cattttaaata 660
aacgtacgat gagcttctct ctccaaatccc cttttttttt ttttttttaa aaaaatattt 720
ttttttatata caatccaaat gggctccaaat ttatcataaa ttaggttagaa acttagatata 780
taaagaaaaga aaagggttta tctcgcaagt gtggctatgg tgggacgtgt caaattttgg 840
attgttagcca aacatgagat ttgatttaaa gggattggc caaatcaccg aaagcaggca 900

| | |
|---|------|
| tcttcatcat aaatttagttt gtttatattt acagaattat acgctttac tagttatagc | 960 |
| atccggtatac ttttctggg taactgccaa accaccacaa atttcaagtt tccatttaac | 1020 |
| tcttcaactt caacccaacc aaatttattt gcttaattgt gcagaaccac tccctatatc | 1080 |
| ttcttaggtgc ttccatttgt tccgaggtaa gaaaagattt ttgtttcttt gaatgcttta | 1140 |
| tgccactcgt ttaacttctg aggtttgtgg atcttttagg cgactttttt ttttttgtta | 1200 |
| tgtaaaattt gtttcataaa tgcttctcaa cataaatctt gacaaagaga aggaatttttta | 1260 |
| ccaagtattt aggttcagaa atggataatt ttcttactgt gaaatatcct tatggcaggt | 1320 |
| tttactgtta tttttcactgtt aa | 1342 |

<210> 9

<211> 1008

<212> DNA

<213> *Lycopersicon esculentum*

220

<221> promoter

<222> (1) .. (1008)

<223> PDS Promotor

<400> 9
aagcttgacg atttccttc agagttagaaa ttgaaaaagaa tcaactaaaa aggatagtcc 60
ttcgatttga tttccggctt aaaaataaaac taataagaat gagagagcga ataatagaat 120
atttgaaat tttaaagata ttcaactatg ttaaattgcg ttataaattt cttaaattag 180
tagcacctaa tagtttagtt ctcaaaagtc aaaaactacta cataatcgtg ctcattttc 240
acattaaaat gcctacatga tgtaaaagta aaactcgttag cattctacgt gtttactca 300
actcaaacat cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt 360
tctttttttt ttttaaaaaa atattttttt ttatatcaat ccaaatgggc tccaatttat 420
cataaaattag gtagaaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg 480
ctatggtggg acgtgtcaaa ttttggattt tagccaaaca tgagatttga tttaaaggaa 540

atggccaaa tcaccgaaag caggcatctt catcataaat tagttgttt atttatacag 600
aattatacgc ttttactagt tatagcattc ggtatcttt tctggtaac tgccaaacca 660
ccacaaaattt caagttcca tttaactctt caacttcaac ccaaccaaattt ttatggctt 720
aattgtgcag aaccactccc tataatcttct aggtgcttcc attcggtccg aggtaagaaa 780
agattttgtt ttctttgaat gctttatgcc actcggtttaa cttctgaggt ttgtggatct 840
tttaggcac tttttttttt tttgtatgta aaatggattt cataaatgct tctcaacata 900
aatcttgaca aagagaagga atttaccaa gtatggatgt tcagaaatgg ataatttct 960
tactgtgaaa ttccttatg gcagggttta ctgttatttt tcagtaaa 1008

<210> 10

<211> 2078

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1) .. (2078)

<223>

<400> 10 ;
tttgcagta ttacaacagc ttatatgtt agcaggtaaa agcttcaatg ccctattctt 60
tctacagtta tcaatgttgc tcgtctaata tctgggttgc ttctcgaaat gtcaattggc 120
ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgc 180
atttatataat ttcatagttt tcagtctcaa ttcttttagct gtttcctcat agtaaagttg 240
tctaataatga aatgaaaatg ttctgtgtgt tgtactaata ccttttcattg gttgtctata 300
gaacgtcgat gaagagccaa acagaaaacta ttttgggtcg cgatttctga taccattgt 360
tctgaatgct gggtgggagc tcatacagaag ctttacaatg ggtcacatat atggagccga 420
gtatgaggaa tgctggaaat cagttgtgct tcgcgtgcta ggactttcc ttcctggat 480
ttctgcccac agcccagttt attacgtgaa ctccgtcaga cttggaaagg agagaagtac 540

ccaaatgtcg tcTTTTtaga aatactttg tcacaaaata gcggggTTTA cagctacaga 600
agatcatgca gaaggcgtcc agtttagttt ttGAAGGTTG tttGGAGTTT atttatctaa 660
agtaaaactta aatcagcttt ttGTTTATGA GTTCAGTGAA ctatATGTTC aaATAAGACT 720
tccCTTGTa gaatatgtgt tttttttgt tGTTGAGCAC tttGTGTGCA ttggataaac 780
ccccaaCgtg taatAGCTAC catacaAGAG aAGTAACtCG cACTGTCCAT GtCTTATGTG 840
GCTCGACTCA gaaAGCATTc aggGGGATTG ataACCACCC TCCAAACCAA CTGAACCATT 900
gtGAATAACC ACCCTTCAAA tcaaccGAGt CCTCGTGAAG GACAATATG TGGTTTATA 960
tacattaaat ttGTTTTTA catGCTTCCT CTTACTTCTT tagTTTCTT gaccatATCT 1020
tCTTTTCCC ttCTGTaATT gacATTTCT tcaaACCATC cAGCAATGTG gaAGCTTyAC 1080
gATTTCTT CAGAGTAGAA ATTGAAAAGA ATCAACTAAA AaggatAGTC CTTGATTG 1140
atTTCCGGCT taaaaataaa ctaataAGAA tgAGAGAGCG aataATAGAA tATTTGAAA 1200
ttttaaAGAT attcaactat gttaaATTGC gttataaATT tcttaaATTa gtagcaccta 1260
atAGTTAGT tctAAAAGT caAAactACT acataATGTG CTCATTTTC acattAAAAT 1320
gcctacatga tGtAAAAGTA AAACtCGTAG CATTCTACGT GTTTACTCA ACTCAAACAT 1380
cctgttcatt ttaataAAACG tacGATGAGC ttctctctcc aATTTCTT tCTTTTTT 1440
ttttAAAAAA atATTTTTT ttatATCAAT cAAATGGGC TCCAAATTAT cataAAATTAG 1500
gtAGAAACTT agatATTAAGA gaaAGAAAAG gTTTATCTC GCAAGTGTGG CTATGGTGG 1560
acgtgtcaaa ttttgatttG tagccAAACA tgAGATTGAt tttAAAGGGA attggccAA 1620
tcaccgaaAG caggcatCTT catcataAAAT tagTTGTTT atttatacAG aattatacGC 1680
ttttactagt tatAGCATTc ggtatTTTT tctGGGTaAC tgccAAACCA ccacAAATT 1740
caagTTTCCA ttAAACTCTT caacttcaAC ccaacAAAT ttATTTGCTT aattgtgcAG 1800
aaccactccc tatATCTTCT aggtGCTTC attcgTTCCG aggtAAAGAA agATTTTGT 1860
ttctttGAAT gCTTTATGCC actcgTTAA CTTCTGAGGT ttgtggatCT tttaggcGAC 1920
tttttttttt tttgtatGTA aaATTTGTTT cataAAATGCT tctcaACATA aatCTTGACA 1980
aAGAGAAGGA attttaccaa GTATTTAGGT tcAGAAATGG ataATTTCT tactGTGAAA 2040
tatCCTTATG gcaggTTTA ctgttatttt tcagtaaa 2078

<211> 1528

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)...(1528)

<223> CHRC Promotor

<400> 11

| | |
|--|------|
| tacaaaattag ggtaacttta ttcattttca tccattctct ttattgttaa attttgtaca | 60 |
| tttattcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca | 120 |
| ctcaaggcctt taccatcttc cttttctatt tcaatactat ttctacttca tttttcacgt | 180 |
| ttttaacatc tttctttatt tcttgccac ttcgtttagg gatgcctaatt gtcccaaatt | 240 |
| tcatctctcg tagtaacaca aaaccaatgt aatgctactt ctctctacat ttttaataca | 300 |
| aataaaagtga aacaaaatat ctataaataa acaaataatat atatttgtt agacgctgtc | 360 |
| tcaacccatc aattaaaaaa ttttgttata tttctacttt acctactaaa tttgtttctc | 420 |
| atatttacct ttaaaccccc aaaaaaaaa attataaaaa agaaagaaaa aagctaaacc | 480 |
| ctatttaaat agctaactat aagatcttaa aattatcctc atcagtgtat agtttaattg | 540 |
| gttattaact tataacatta tatatctatg acatatactc ttccttagct atttctcaca | 600 |
| tttttaact taagaaaata gtcataacat agtctaaaat tcaaaacatcc acatgctcta | 660 |
| attgattaa caaaaagtta gaaatattta tttaataaaa aaagactaat aaatataaa | 720 |
| aatgaatgtt catacgcaga cccattttaga gatgagtatg cttcacatg ctgagattat | 780 |
| tttcaaaaact aagggtttag caatattaaa tcaataaaaat tattataat aacaaaatta | 840 |
| acctgctcgt gtttgctgta tatgggaggc tacaaaataa attaaactaa agatgattat | 900 |
| gttttagaca tttttctat ctgtattagt ttatacatat taattcagga gctgcacaac | 960 |
| ccaattctat ttccgttcct tggggctgg gtttctcaca aggtcaata gtcaatatta | 1020 |
| ggttttattg gacttttaat agtatcaaac aaatctatgt gtgaacttaa aaattgtatt | 1080 |
| aaatatttag ggtaacctgt tgccgtttt agaataatgt ttcttcttaa tacacgaaag | 1140 |

cgtattgtgt attcattcat ttggcgcc tc acatgcttcg gttggctgc ttttagtctct 1200
gccttctttg tatattgtac tccccctctt cctatgccac gtgttctgag cttaacaagc 1260
cacgttgcgt gccattgcca aacaagtcat tttaacttca caaggccga tttgacctcc 1320
aaaacaacga caagttccg aacagtgcgc aagatcaagg gtataatcgt ctttttgaat 1380
tctatttctc ttatattaa agtccctctc gtgtgatagt ttttaaaaga tttttaaaac 1440
gtagctgctg tttaagtaaa tcccagtcct tcagtttg ttttgtgtg ttttgttct 1500
ctgatttacg gaatttggaa ataattct 1528

<210> 12

<211> 1538

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)..(1538)

<223> CHRC Promotor

<400> 12

tacaaatttag gtttacttta ttcatttca tccattctct ttattgttaa attttgtaca 60
tttattcaat aatatttatat gtttattaca aattctcact ttcttattca tacctattca 120
ctcaaggcctt taccatcttc cttttctatt tcaatactat ttctacttca ttttcacgt 180
ttttaacatc tttctttatt tcttgcac ttcgtttagg gatgccta at gtc cca aatt 240
tcatctctcg tagtaacaca aaaccaatgt aatgctactt ctctctacat tttaatacaca 300
aataaaagtga aacaaaatat ctataaataa acaaataatat atatttgtt agacgctgtc 360
tcaacccatc aattaaaaaa tttgttata tttctacttt acctactaaa tttgttctc 420
atatttacct tttaaccccc aaaaaaaaaa attataaaaa agaaagaaaa aagctaaacc 480
ctattnaat agctaactat aagatcttaa aattatccctc atcagtgtat agtttaatg 540
gttattaact tataacatta tataatctatg acatataactc tctcctagct atttctcaca 600

tttttaact taagaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta 660
attgattaa caaaaagtta gaaatattta tttaaataaa aaagactaat aaatatataa 720
aatgaatgtt catacgcaga cccatttaga gatgagtagt cttcacatg ctgagattat 780
tttcaaaact aaggttgtag caatattaaa tcaataaaat tattataaaat aacaaaatta 840
acctgctcg gtggctgta tatgggaggc tacaaaataa attaaactaa agatgattat 900
gttttagaca tttttctat ctgtatttagt ttatacatat taattcagga gctgcacaac 960
ccaattctat ttgcgttcct tggtggctgg gtttctcaca aggtcaata gtcataat 1020
ggtttattg gacttttaat agtatcaaac aaatctatgt gtgaacttaa aaattgtatt 1080
aaatatttag ggtaacctgt tgccgtttt agaataatgt ttcttcttaa tacacgaaag 1140
cgtattgtgt attcattcat ttggcgccctc acatgcttcg gttggctcgc ttttgtct 1200
gccttcttg tattttgtac tccccctctt cctatgccac gtgttctgag cttaacaagc 1260
cacgttgcgt gccattgcca aacaagtcat tttaacttca caaggccga tttgacctcc 1320
aaaacaacga caagttccg aacagtgcgc aagatcaagg gtataatcgt cttttgaat 1380
tctatttctc ttatattaaat agtcctctc gtgtgatagt tttttaaaga tttttaaaac 1440
gtatgtctg tttaagtaaa tcccagtcct tcagtttgcg cttttgtgtg ttttgttct 1500
ctgatttacg gaatttggaa ataatttctc taaaggc 1538

<210> 13

<211> 1525

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)..(1525)

<223> CHRC Promotor

<400> 13

tacaatttag ggttacttta ttcatattca tccattctct ttattgttaa attttgtaca 60

| | |
|---|------|
| tttattcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca | 120 |
| ctcaaggcctt taccatcttc cttttctatt tcaatactat ttctacttca ttttcacgt | 180 |
| tttaaacatc tttctttatt tcttgccac ttcgtnnagg gatgccta at gtcggaaatt | 240 |
| tcatctctcg tagtaacaca aaaccaatgt aatgctacct ctctctacat tttaatacaca | 300 |
| aataaagtga aacaaaatat ctataaataa acaaataatat atattnnntt agacgctg | 360 |
| tcaacccatc aattaaaaaa ttttgtata ttctacttt acctactaaa tttgtttctc | 420 |
| atatttacct tttaaccccc aaaaaaaaa attataaaaa agaaagaaaa aagctaaacc | 480 |
| ctatttaaat agctagctat aagatcttaa aattatccctc atcagtgtat agtttaattg | 540 |
| gttattaact tataacat tatactatg acatatactc tctcttagt atttctcaca | 600 |
| tttttttaact taagaaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta | 660 |
| atttgattaa caaaaagata gaaatattta tttaaataaa aaagactaat aaatataaa | 720 |
| aatgaatgtt catacgcaga cccattnaga gatgagttatg ctttcacatg ttgagattat | 780 |
| tttcaaaaact aagggtttag caatattaaa tcaataaaaat tattataat aacaaaatta | 840 |
| acctgctcggtt gtttgctgaa tatgggaggc tacaaaataa attaaactaa agatgattat | 900 |
| gttttagaca tttttctat ctgtattatg ttatacatat taattcagga gctgcacaac | 960 |
| ccaaattctat ttctggttctt tggtggctgg gtttctcaca aggtcaata gtcaatatta | 1020 |
| ggttttattt gacttttaat agtataat aaatctatgt gtgaacttaa aaattgtatt | 1080 |
| aaatattttat ggttaacctgt tgccgtttt agaataatgt ttcttcttaa tacacgaaag | 1140 |
| cgtattgtgtt attcattcat ttggcgccctc acatgcttcg gttggctcgc tttagtctct | 1200 |
| gccttctttt gatttgtac tccccctctt cctatgccac gtgttctgag cttaacaagc | 1260 |
| cacgttgcgt gccattgcca aacaagtcat tttaacttca caaggtccga tttgacctcc | 1320 |
| aaaacaacga caagttccg aacagtcgag aagatcaagg gtataatcgt ctttttgaat | 1380 |
| tctatttctc ttatatttaat agtccctctc gtgtgatagt tttaaaaga tttaaaaac | 1440 |
| gtagctgctg tttaagtaaa tcccagtcct tcagtttgcg cttttgcgtg ttgtttct | 1500 |
| ctgatttacg gaatttggaa ataat | 1525 |

<210> 14

<211> 1519

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)..(1519)

<223> CHRC Promotor

| | |
|---|------|
| <400> 14 | |
| ttactttattt cattttcatc cattctcttt attgttaaat tttgtacatt tattcaataa | 60 |
| tattatatgt ttattacaaa ttctcacttt ctattcata cctattcaact caagccttta | 120 |
| ccatcttcctt tttcttatttc aatactatttt ctacttcattt tttcacgttt ttaacatctt | 180 |
| tctttatttc ttgtccactt cgtttaggaa tgccataatgtt cccaaatttc atctctcgta | 240 |
| gtaacacaaaa accaatgtaa tgctacctctt ctctacatttt ttaatacaaa taaagtgaaa | 300 |
| caaaaatatct ataaataaaac aaatataat attttggtag acgcgtgtc aacccatcaa | 360 |
| ttaaaaaattt ttgttatattt tctactttac ctactaaattt tgtttctcatt atttaccttt | 420 |
| taacccccac aaaaaaaaaat tataaaaaag aaagaaaaaa gctaaaccctt atttaatag | 480 |
| ctagctataa gatcttaaaa ttatcctcat cagtgtatag tttaattgggt tattaactta | 540 |
| taacattata tatctatgac atatactctc tccttagctat ttctcacattt ttttaactta | 600 |
| agaaaaatgtt cataacatag tctaaaatttc aaacatccac atgccttaat ttgattaaca | 660 |
| aaaagataga aatattttttaaataaaaaa agactaataa atatataaaaa tgaatgttca | 720 |
| tacgcagacc catttagaga tgagtatgctt ttcacatgtt gagattttt tcaaaaactaa | 780 |
| ggttttagca atattaaatc aataaaatattt ttataaataa caaaattaac ctgctcggt | 840 |
| ttgctgaata tggaggccta caaaaataat taaactaaag atgattatgtt tttagacattt | 900 |
| ttttctatct gtatttagttt atacatatta attcaggagc tgcacaaccc aattctat | 960 |
| tcgttccttg gtggctgtgt ttctcacaag gttcaatagt caatattagg ttttattgga | 1020 |
| cttttaatag tatcaaataa atctatgtgt gaactaaaa attgtattaa atattttaggg | 1080 |
| taacctgttg ccgttttag aataatgtttt cttcttaata cacgaaagcg tattgtgtat | 1140 |
| tcattcattt ggccctcac atgcctcggt tggctcgctt tagtctctgc cttctttgtat | 1200 |

ttttgtactc cccctttcc tatgccacgt gttctgagct taacaagcca cgttgcgtgc 1260
 cattGCCaaa caagtcattt taacttcaca aggtccgatt tgacctccaa aacaacgaca 1320
 agtttccgaa cagtcgcgaa gatcaagggt gtaatcgct ttttgaattc tatttcttt 1380
 tatttaatag tccctcttgt gtgatagttt tttaaagatt tttaaaacgt agctgctgtt 1440
 taagtaaattc ccagtcccttc agtttgct tttgtgtgtt ttgtttctct gatttacgga 1500
 atttggaaat aataagctt 1519

<210> 15

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<223>

<400> 15

ggcacgagct tgcacgcaag tcagcgcgca agtcaca cctggcggtc cacagcctca 60
 aataataaag agtcagaagcg tttgtgcgcc tcgacgtggc cagtcgtcac tgccttgaac 120
 ccgcgagtct cccgcgcac tgactgccc agcacagcta gacga atg cag cta gca 177
 Met Gln Leu Ala
 1

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
 Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
 5 10 15 20

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
 Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
 25 30 35

gcg acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
 Ala Thr Gln Tyr Ser Leu Pro Ser Glu Ser Asp Ala Ala Arg Pro
 40 45 50

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369

| | | |
|---|-----|------|
| Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile | | |
| 55 | 60 | 65 |
| aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac | | 417 |
| Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His | | |
| 70 | 75 | 80 |
| gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg | | 465 |
| Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp | | |
| 85 | 90 | 95 |
| ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc | | 513 |
| Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser | | |
| 105 | 110 | 115 |
| ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca | | 561 |
| Leu Leu Asp Ile Val Val Phe Val Leu Glu Phe Leu Tyr Thr | | |
| 120 | 125 | 130 |
| ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg | | 609 |
| Gly Leu Phe Ile Thr His Asp Ala Met His Gly Thr Ile Ala Met | | |
| 135 | 140 | 145 |
| aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg | | 657 |
| Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu | | |
| 150 | 155 | 160 |
| tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac | | 705 |
| Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His | | |
| 165 | 170 | 175 |
| cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga | | 753 |
| His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly | | |
| 185 | 190 | 195 |
| aac cct ggc att gtg ccc tgg ttt gcc agc ttc atg tcc agc tac atg | | 801 |
| Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met | | |
| 200 | 205 | 210 |
| tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag | | 849 |
| Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln | | |
| 215 | 220 | 225 |
| ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg | | 897 |
| Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala | | |
| 230 | 235 | 240 |
| ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ccc | | 945 |
| Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro | | |
| 245 | 250 | 255 |
| cac aag cct gag cct ggc gcc gcg tca ggc tct tca cca gcc gtc atg | | 993 |
| His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met | | |
| 265 | 270 | 275 |
| aac tgg tgg aag tcg cgc act agc cag gcg tcc gac ctg gtc agc ttt | | 1041 |

| | | |
|--|-----|------|
| Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe | | |
| 280 | 285 | 290 |
| ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc | | 1089 |
| Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro | | |
| 295 | 300 | 305 |
| ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga | | 1137 |
| Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg | | |
| 310 | 315 | 320 |
| ggt ctg gtt cct gcc tag ctggacacac tgcagtggc cctgctgcc | | 1185 |
| Gly Leu Val Pro Ala | | |
| 325 | | |
| gctggcatg caggttgtgg caggactgg tgaggtgaaa agctgcaggc gctgctgcc | | 1245 |
| gacacgctgc atggctacc ctgtgttagct gcugccacta ggggaggggg tttgttagctg | | 1305 |
| tcgagcttgc cccatggatg aagctgtgta gtggcagg gactacaccc acaggccaac | | 1365 |
| acccttgcag gagatgtctt gcgtcggag gagtggtgg cagtgttagat gctatgattg | | 1425 |
| tatcttaatg ctgaagcctt tagggagcg acacttagtgc tggcaggc aacgcctgc | | 1485 |
| aaggtgcagg cacaagctag gctggacgag gactcggtgg caggcaggta aagaggtgc | | 1545 |
| ggagggtggt gccacaccca ctggcaaga ccatgctgca atgctggcgg tgtggcagt | | 1605 |
| agagctgcgt gattaactgg gctatggatt gttttagcag tctacttat tctttgatat | | 1665 |
| agataactggt caggcaggc aggagagtga gtatgaacaa gttgagaggt ggtgcgtgc | | 1725 |
| ccctgcgtt atgaagctgt aacaataaag tggttcaaaa aaaaaa | | 1771 |

<210> 16

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 16

| | | |
|---|---|----|
| Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala | | |
| 1 | 5 | 10 |
| 15 | | |

| | | |
|---|----|----|
| Glu Ala Leu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val | | |
| 20 | 25 | 30 |

22

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
325

<210> 17

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168)..(1130)

<223>

<400> 17
cggggcaact caagaaattc aacagctgca agcgccgcggccccc agcctcacag cgccaaagtga 60
gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg 120
ctccgtcctc tgccaaatct cgcgtcgaaaa cctgcctaag tcgaaga atg cac gtc 176
Met His Val
1
gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224
Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
5 10 15
agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272
Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
20 25 30 35

| | | | | |
|--|-----|-----|-----|-----|
| gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro | 40 | 45 | 50 | 320 |
| cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly | 55 | 60 | 65 | 368 |
| acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro | 70 | 75 | 80 | 416 |
| aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala | 85 | 90 | 95 | 464 |
| cag ctt ttg ggc gga agc agc cta ctg cac atc gct gca gtc ttc Gln Leu Leu Gly Gly Ser Ser Leu Leu His Ile Ala Ala Val Phe | 100 | 105 | 110 | 512 |
| att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp | 120 | 125 | 130 | 560 |
| gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu | 135 | 140 | 145 | 608 |
| ctt ggc aac atc tgc ata tca ctg tac gac tac agc atg Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met | 150 | 155 | 160 | 656 |
| ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly | 165 | 170 | 175 | 704 |
| aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe | 180 | 185 | 190 | 752 |
| gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu | 200 | 205 | 210 | 800 |
| gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn | 215 | 220 | 225 | 848 |
| ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc Leu Leu Val Phe Met Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu | 230 | 235 | 240 | 896 |
| ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala | 245 | 250 | 255 | 944 |

| | |
|---|------|
| gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca | 992 |
| Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala | |
| 260 265 270 275 | |
| tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg | 1040 |
| Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp | |
| 280 . 285 290 | |
| gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc | 1088 |
| Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys | |
| 295 . . 300 . . 305 | |
| cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga | 1130 |
| Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala | |
| 310 315 320 | |
| cctggtccct ccgctggta cccagcgct gcacaagagt gtcatgctac agggtgctgc | 1190 |
| ggccagtgcc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca | 1250 |
| ctgggcatgc cactgagcac tggcgctgct actgagcaat gggcgctgcta ctgagcaatg | 1310 |
| ggcgtgctac tgacaatggg cgtgctactg gggctggca gtggcttagga tggagtttga | 1370 |
| tgcattcagt agcggtggcc aacgtcatgt ggatggtgaa agtgctgagg ggtttaggca | 1430 |
| gccggcattt gagagggcta agttataaat cgcatgctgc tcatgcgcac atatctgcac | 1490 |
| acagccaggg aaatcccttc gagagtgatt atgggacact tgtattggtt tcgtgctatt | 1550 |
| gttttattca gcagcagtac ttagtgaggg tgagagcagg gtggtgagag tggagtgagt | 1610 |
| gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct | 1662 |

<210> 18

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 18

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95 //

Ala Thr Ala Gln Leu Leu Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

//

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

// : /

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

: /

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val
180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
195 200 205

:

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro
210 215 220

:

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala
225 230 235 240

:

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro
245 250 255

27

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr
260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp
275 280 . 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu
290 295 300

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | His | Cys | Arg | Arg | Leu | Ser | Gly | Arg | Gly | Leu | Val | Pro | Ala | Leu | Ala |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |

<210> 19

<211> 729

<212> DNA

<213> *Agrobacterium aurantiacum*

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 19

```

atg agg gca ; cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1           .   .      5                   10                  15

```

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 . 20 25 30

gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 . 60

cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat 240
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 . 75 80

| | |
|---|-----|
| gct gct atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95 | 288 |
| cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr 100 105 110 | 336 |
| gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125 | 384 |
| cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 135 140 | 432 |
| gtc atc gtg acg gtc tat gct ctg atc ctt ggg gat cgc tgg atg tac Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 155 160 | 480 |
| gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gct tcg atc cag ctg ttc Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175 | 528 |
| gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gct ttc ccg Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190 | 576 |
| gac cgc cac aat gct cgg tcg ccg atc agc gac ccc gtg tcg ctg Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205 | 624 |
| ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220 | 672 |
| ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240 | 720 |
| acc gca tga Thr Ala | 729 |

<210> 20

<211> 242

<212> PRT

<213> Agrobacterium aurantiacum

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 21

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99) .. (827)

<223>

<400> 21

ctgcaggccg ggcccggtgg ccaatggtcg caaccggcag gactggaaca ggacggcggg 60

ccggctctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116
Met Ser Gly Arg Lys Pro
1 / 5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164
Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile
10 15 20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212
Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp
25 30 35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260
Ala Ala Ala His Pro Leu Ala Val Leu Cys Leu Ala Gly Leu Thr
40 45 50

tgg ctg tcg gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308
Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly
55 60 65 70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356
Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu
75 80 85

| | |
|---|-----|
| gct ctg tgg ctc tat gct ggg ttc tcg tgg ccc aag ctg atc gcc aag Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys 90 95 100 | 404 |
| cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe 105 110 115 | 452 |
| ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr 120 125 130 | 500 |
| ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat Phe Gly Trp Arg Glu Gly Leu Leu Pro Val Ile Val Thr Thr Tyr 135 140 145 150 | 548 |
| gct ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val 155 160 165 | 596 |
| ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu 170 175 180 | 644 |
| ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gct agg Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg 185 190 195 | 692 |
| tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Thr Cys Phe His Phe 200 205 210 | 740 |
| ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp 215 220 225 230 | 788 |
| cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala 235 240 | 837 |
| cattgtcggt gcgacagtcc tcgtgatggc gctgaccgcc tattccgtcc accgctggat 957 | 897 |
| tatgcacggc cccctaggct gggctggca caagtccccat cacgaagagc acgaccacgc 1017 | |
| gttggagaag aacgacacct acggcgctcg tttcgccgtg ctggcgacga tcctttcac 1077 | |
| cgtggcgcc tattgggtggc cggtgctgtg gtggatcgcc ctggcatga cggctatgg 1137 | |
| tcccgccgg ggctatttcc gcaggctcta ccaagctcat cgcctgcacc acgcggcga 1197 | |
| ggggcgggac cactgcgtca gcttcggctt catctatgcc ccacccgtgg acaagctgaa 1257 | |
| gcaggatctg aagcggtcgg gtgtcctgcg cccccaggac gagcgtccgt cgtgatctct 1317 | |

gatccggcg tggccgcgt aaatccgacg tgctgctggc aggggcccgc cttgccaacg 1377
gactgatcgc gctggcgatc cgcaaggcgc ggcccgacct tcgcgtgctg ctgctggacc 1437
gtgcggcggg cgccctcgac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc 1497
actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcgg 1557
tcccagacca ttgcgaaagg ctccggccg gatatggctc gatcgacggg cgggggctga 1617
tgcggtgaccc gacc 1631

<210> 22

<211> 242

<212> PRT

<213> Alcaligenes sp.

<400> 22

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe
20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala His Pro Leu Leu Ala Val Leu
35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
225 230 235 240

Arg Ala

<210> 23

<211> 729

<212> DNA

<213> Paracoccus marcusii

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 23

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg

| | | | |
|---|-----|-----|-----|
| Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu | | | |
| 1 | 5 | 10 | 15 |
| atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat | | | 96 |
| Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His | | | |
| 20 | 25 | 30 | |
| gcg ctg tgg ttt ctg gac gcg ggc cat ccc atc ctg gcg gtc gcg | | | 144 |
| Ala Leu Trp Phe Leu Asp Ala Ala His Pro Ile Leu Ala Val Ala | | | |
| 35 | 40 | 45 | |
| aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg | | | 192 |
| Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala | | | |
| 50 | 55 | 60 | |
| cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat | | | 240 |
| His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn | | | |
| 65 | 70 | 75 | 80 |
| gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg | | | 288 |
| Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp | | | |
| 85 | 90 | 95 | |
| cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc | | | 336 |
| Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr | | | |
| 100 | 105 | 110 | |
| gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc | | | 384 |
| Asp Asp Asp Pro Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala | | | |
| 115 | 120 | 125 | |
| cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc | | | 432 |
| Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro | | | |
| 130 | 135 | 140 | |
| gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac | | | 480 |
| Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr | | | |
| 145 | 150 | 155 | 160 |
| gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc | | | 528 |
| Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe | | | |
| 165 | 170 | 175 | |
| gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg | | | 576 |
| Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro | | | |
| 180 | 185 | 190 | |
| gac cgc cat aat gcg ccg tcg tcg ccg atc agc gac cct gtg tcg ctg | | | 624 |
| Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu | | | |
| 195 | 200 | 205 | |
| ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac | | | 672 |
| Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His | | | |
| 210 | 215 | 220 | |
| ccg acg gtg ccg tgg tgg ccg ctg ccc agc acc cgc acc aag ggg gac | | | 720 |

35

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

acc gca tga 729
Thr Ala

<210> 24

<211> 242

<212> PRT

<213> Paracoccus marcusii

<400> 24

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 25

<211> 1629

<212> DNA

<213> Synechocystis

<220>

<221> CDS

<222> (1)..(1629)

<223>

<400> 25

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
1 5 10 15

gtc tgt gca gcc tat ttg ctc caa cg ggc ttg ggg gtg acg tta cta 96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
20 25 30

| | |
|---|-----|
| gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met 35 40 45 | 144 |
| ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His 50 55 60 | 192 |
| gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln 65 70 75 80 | 240 |
| tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly 85 90 95 | 288 |
| ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys 100 105 110 | 336 |
| gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln 115 120 125 | 384 |
| ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe 130 135 140 | 432 |
| aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp 145 150 155 160 | 480 |
| gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala 165 170 175 | 528 |
| ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn 180 185 190 | 576 |
| gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys 195 200 205 | 624 |
| tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met 210 215 220 | 672 |
| atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly 225 230 235 240 | 720 |
| ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln 245 250 255 | 768 |

| | | |
|---|-----|------|
| ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa | | 816 |
| Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu | | |
| 260 | 265 | 270 |
| aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg | | 864 |
| Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg | | |
| 275 | 280 | 285 |
| gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg | | 912 |
| Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu | | |
| 290 | 295 | 300 |
| caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg | | 960 |
| Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly | | |
| 305 | 310 | 315 |
| 320 | | |
| gaa cga ctg gaa ccg cgc act gtg aac aat aac gaa gcc att tta aaa | | 1008 |
| Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys | | |
| 325 | 330 | 335 |
| atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg | | 1056 |
| Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly | | |
| 340 | 345 | 350 |
| ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat | | 1104 |
| Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His | | |
| 355 | 360 | 365 |
| gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct | | 1152 |
| Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala | | |
| 370 | 375 | 380 |
| aat ccg tct ttat ttg gat att ccc act gta ttg gac ccc acc atg | | 1200 |
| Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met | | |
| 385 | 390 | 395 |
| 400 | | |
| gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttg gcc ccc tac | | 1248 |
| Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr | | |
| 405 | 410 | 415 |
| cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc | | 1296 |
| Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr | | |
| 420 | 425 | 430 |
| gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg | | 1344 |
| Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr | | |
| 435 | 440 | 445 |
| gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa | | 1392 |
| Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu | | |
| 450 | 455 | 460 |
| agt ccc gcc gaa ctg gcc caa ccg ctg gga agt tac aac ggc aat gtc | | 1440 |
| Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val | | |
| 465 | 470 | 475 |
| 480 | | |

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
 485 490 495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
 530 535 540

<210> 26

<211> 542

<212> PRT

<213> Synechococystis

<400> 26

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
 50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
 65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
 85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
 100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met
210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly
225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu
290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly
340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His
355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala
370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met
385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 535 540

<211> 776

<212> DNA

<213> *Bradyrhizobium* sp.

<220>

<221> CDS

<222> (1)..(774)

<223>

<400> 27

| | |
|---|----|
| atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc | 48 |
| Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg | |
| 1 5 10 15 | |

| | |
|---|----|
| gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc | 96 |
| Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile | |
| 20 25 30 | |

| | |
|---|-----|
| atc gcc gcc tgg ctg gtg cat gtc ggt ctg atg ttc ttc tgg ccg | 144 |
| Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro | |
| 35 40 45 | |

| | |
|---|-----|
| ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag | 192 |
| Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln | |
| 50 55 60 | |

| | |
|---|-----|
| acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac | 240 |
| Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His | |
| 65 70 75 80 | |

| | |
|---|-----|
| ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag | 288 |
| Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln | |
| 85 90 95 | |

| | |
|---|-----|
| ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc | 336 |
| Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val | |
| 100 105 110 | |

| | |
|---|-----|
| gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat | 384 |
| Glu His His Lys His Arg His Pro Gly Thr Ala Glu Asp Pro Asp | |
| 115 120 125 | |

| | |
|---|-----|
| ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt | 432 |
| Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe | |
| 130 135 140 | |

| | |
|---|-----|
| ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc | 480 |
|---|-----|

| | | | |
|---|-----|-----|-----|
| Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val | | | |
| 145 | 150 | 155 | 160 |
| | | | |
| tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg | | | 528 |
| Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu | | | |
| 165 | 170 | 175 | |
| | | | |
| ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc | | | 576 |
| Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr | | | |
| 180 | 185 | 190 | |
| | | | |
| ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat | | | 624 |
| Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp | | | |
| 195 | 200 | 205 | |
| | | | |
| cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg | | | 672 |
| Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu | | | |
| 210 | 215 | 220 | |
| | | | |
| acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat | | | 720 |
| Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp | | | |
| 225 | 230 | 235 | 240 |
| | | | |
| gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg | | | 768 |
| Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg | | | |
| 245 | 250 | 255 | |
| | | | |
| cgt gac ta | | | 776 |
| Arg Asp | | | |

<210> 28

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 28

| | | | |
|---|---|----|----|
| Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg | | | |
| 1 | 5 | 10 | 15 |

| | | |
|---|----|----|
| Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile | | |
| 20 | 25 | 30 |

| | | |
|---|----|----|
| Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro | | |
| 35 | 40 | 45 |

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
145 150 155 160

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
245 250 255

Arg Asp

<210> 29
<211> 777
<212> DNA
<213> Nostoc sp.

<220>

<221> CDS
<222> (1)..(777)
<223>

| | | |
|---|--|-----|
| <400> 29 | | |
| atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta | | 48 |
| Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu | | |
| 1 5 10 15 | | |
| ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt | | 96 |
| Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe | | |
| 20 25 30 | | |
| att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta | | 144 |
| Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu | | |
| 35 40 45 | | |
| ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc | | 192 |
| Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala | | |
| 50 55 60 | | |
| atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat | | 240 |
| Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His | | |
| 65 70 75 80 | | |
| gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat | | 288 |
| Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn | | |
| 85 90 95 | | |
| ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa | | 336 |
| Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys | | |
| 100 105 110 | | |
| gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat | | 384 |
| Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp | | |
| 115 120 125 | | |
| tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg | | 432 |
| Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp | | |
| 130 135 140 | | |

-211- 258

5213- BBT

<213> Nostoc sp.

<400> 30

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Gln | Cys | Gln | Pro | Ser | Ser | Leu | His | Ser | Glu | Lys | Leu | Val | Leu |
| 1 | | | | | | | | | 10 | | | | | | 15 |

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
245 250 255

Ser Leu

<210> 31
<211> 831
<212> DNA
<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (1)..(831)

<223>

| | | |
|---|--|-----|
| <400> 31 | | |
| atg cca tcc gag tcg tca gac gca gct cgt cct gtg ttg aag cac gcc | | 48 |
| Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Val Leu Lys His Ala | | |
| 1 5 10 15 | | |
| tat aaa cct cca gca tct gac gcc aag ggc atc act atg gcg ctg acc | | 96 |
| Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr | | |
| 20 25 30 | | |
| atc att ggc acc tgg acc gca gtg ttt tta cac gca ata ttc caa atc | | 144 |
| Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile | | |
| 35 40 45 | | |
| agg cta ccg aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa | | 192 |
| Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu | | |
| 50 55 60 | | |
| gcc aca gcc cag ctg ttg ggc gga agc agc agc cta ttg cac atc gcc | | 240 |
| Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala | | |
| 65 70 75 80 | | |
| gca gtc ttc att gta ctt gag ttt ctg tac act ggt cta ttc atc acc | | 288 |
| Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr | | |
| 85 90 95 | | |
| acg cat gat gca atg cat ggc acc ata gct ttg agg aac agg cag ctc | | 336 |
| Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg Asn Arg Gln Leu | | |
| 100 105 110 | | |
| aat gat ctc ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac | | 384 |
| Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp | | |
| 115 120 125 | | |
| tac agc atg cac tgg gag cac cac aac cat act ggc gaa gtg ggg aaa | | 432 |

| | | | |
|---|-----|-----|-----|
| Tyr Ser Met His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys | | | |
| 130 | 135 | 140 | |
| gac cct gac ttc cac aaa gga aat cct ggc ctt gtc ccc tgg ttc gcc | | | 480 |
| Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe Ala | | | |
| 145 | 150 | 155 | 160 |
| agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cggtt gca | | | 528 |
| Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu Ala | | | |
| 165 | 170 | 175 | |
| tgg tgg gca gtg gtg atg caa acg ttg ggg gcc ccc atg gcg aat ctc | | | 576 |
| Trp Trp Ala Val Val Met Gln Thr Leu Gly Ala Pro Met Ala Asn Leu | | | |
| 180 | 185 | 190 | |
| cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc ttc | | | 624 |
| Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe | | | |
| 195 | 200 | 205 | |
| tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca gca | | | 672 |
| Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala Ala | | | |
| 210 | 215 | 220 | |
| ggc tct cag gtc atg tct tgg ttc agg gcc aag aca agt gag gca tct | | | 720 |
| Gly Ser Gln Val Met Ser Trp Phe Arg Ala Lys Thr Ser Glu Ala Ser | | | |
| 225 | 230 | 235 | 240 |
| gat gtg atg agc ttc ctg aca tgc tac cac ttt gac ctg ttt gcc ccc | | | 768 |
| Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu Phe Ala Pro | | | |
| 245 | 250 | 255 | |
| tgg tgg cag ctg ccc cac tgc cgc cgc ctg tct ggg cgt ggc ctg gtg | | | 816 |
| Trp Trp Gln Leu Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val | | | |
| 260 | 265 | 270 | |
| cct gcc ttg gca tga | | | 831 |
| Pro Ala Leu Ala | | | |
| 275 | | | |

<210> 32

<211> 276

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Val Leu Lys His Ala
1 5 10 15

50

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
20 25 30

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
35 40 45

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
50 55 60

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
65 70 75 80

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
85 90 95

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg Asn Arg Gln Leu
100 105 110

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
115 120 125

Tyr Ser Met His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys
130 135 140

Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe Ala
145 150 155 160

Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu Ala
165 170 175

Trp Trp Ala Val Val Met Gln Thr Leu Gly Ala Pro Met Ala Asn Leu
180 185 190

Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe
195 200 205

Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala Ala
210 215 220

Gly Ser Gln Val Met Ser Trp Phe Arg Ala Lys Thr Ser Glu Ala Ser
225 230 235 240

Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu Phe Ala Pro
 245 250 255

Trp Trp Gln Leu Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val
 260 265 270

Pro Ala Leu Ala
 275

<210> 33

<211> 729

<212> DNA

<213> Paracoccus sp. MBIC1143

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 33

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
 50 55 60

cat gac gcg atg cac ggg tcg gtg ccg ggg cgt ccg cgc gcc aat 240
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 75 80

gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95

| | |
|---|-----|
| cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr 100 105 110 | 336 |
| gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125 | 384 |
| cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 135 140 | 432 |
| gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 155 160 | 480 |
| gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175 | 528 |
| gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190 | 576 |
| gac cgc cac aat gcg ccg tcg ccg atc agc gac ccc gtg tcg ctg Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205 | 624 |
| ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220 | 672 |
| ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240 | 720 |
| acc gca tga Thr Ala | 729 |

<210> 34

<211> 242

<212> PRT

<213> Paracoccus sp. MBIC1143

<400> 34

| |
|--|
| Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 1 5 10 15 |
|--|

Ile Val Ser Gly Gly Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala His Pro Ile Leu Ala Ile Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 35

<211> 735

<212> DNA

<213> Brevundimonas aurantiaca

<220>

<221> CDS

<222> (1) .. (735)

<223>

| | | |
|---|--|-----|
| <400> 35 | | |
| atg acc gcc gcc gtc gcc gag cca cgc acc gtc ccg cgc cag acc tgg | | 48 |
| Met Thr Ala Ala Val Ala Glu Pro Arg Thr Val Pro Arg Gln Thr Trp | | |
| 1 5 10 15 | | |
| atc ggt ctg acc ctg gcg gga atg atc gtg gcg gga tgg gcg gtt ctg | | 96 |
| Ile Gly Leu Thr Leu Ala Gly Met Ile Val Ala Gly Trp Ala Val Leu | | |
| 20 25 30 | | |
| cat gtc tac ggc gtc tat ttt cac cga tgg ggg ccg ttg acc ctg gtg | | 144 |
| His Val Tyr Gly Val Tyr Phe His Arg Trp Gly Pro Leu Thr Leu Val | | |
| 35 40 45 | | |
| atc gcc ccg gcg atc gtg gcg gtc cag acc tgg ttg tcg gtc ggc ctt | | 192 |
| Ile Ala Pro Ala Ile Val Ala Val Gln Thr Trp Leu Ser Val Gly Leu | | |
| 50 55 60 | | |
| tcc atc gtc gcc cat gac gcc atg tac ggc tcc ctg gcg ccg gga cgg | | 240 |
| Phe Ile Val Ala His Asp Ala Met Tyr Gly Ser Leu Ala Pro Gly Arg | | |
| 65 70 75 80 | | |
| ccg cgg ctg aac gcc gca gtc ggc ccg ctg acc ctg ggg ctc tat gcg | | 288 |
| Pro Arg Leu Asn Ala Ala Val Gly Arg Leu Thr Leu Gly Leu Tyr Ala | | |
| 85 90 95 | | |
| ggc ttc cgc ttc gat cgg ctg aag acg gcg cac cac gcc cac cac gcc | | 336 |
| Gly Phe Arg Phe Asp Arg Leu Lys Thr Ala His His Ala His His Ala | | |
| 100 105 110 | | |
| gcg ccc ggc acg gcc gac gac ccg gat ttt cac gcc ccg gcg ccc cgc | | 384 |

| | | | |
|--|-----|-----|-----|
| Ala Pro Gly Thr Ala Asp Asp Pro Asp Phe His Ala Pro Ala Pro Arg | | | |
| 115 | 120 | 125 | |
| gcc ttc ctt ccc tgg ttc ctg aac ttc ttt cgc acc tat ttc ggc tgg Ala Phe Leu Pro Trp Phe Leu Asn Phe Phe Arg Thr Tyr Phe Gly Trp | | | 432 |
| 130 | 135 | 140 | |
| cgc gag atg gcg gtc ctg acc gcc ctg gtc ctg atc gcc ctc ttc ggc Arg Glu Met Ala Val Leu Thr Ala Leu Val Leu Ile Ala Leu Phe Gly | | | 480 |
| 145 | 150 | 155 | 160 |
| ctg ggg gcg cgg gcc aat ctc ctg acc ttc tgg gcc gcg cgg gcc Leu Gly Ala Arg Pro Ala Asn Leu Leu Thr Phe Trp Ala Ala Pro Ala | | | 528 |
| 165 | 170 | 175 | |
| ctg ctt tca gcg ctt cag ctc acc ttc ggc acc tgg ctg cgg cac Leu Leu Ser Ala Leu Gln Leu Phe Thr Phe Gly Thr Trp Leu Pro His | | | 576 |
| 180 | 185 | 190 | |
| cgc cac acc gac cag ccg ttc gcc gac gcg cac cac gcc cgc agc agc Arg His Thr Asp Gln Pro Phe Ala Asp Ala His His Ala Arg Ser Ser | | | 624 |
| 195 | 200 | 205 | |
| ggc tac ggc ccc gtg ctt tcc ctg ctc acc tgt ttc cac ttc ggc cgc Gly Tyr Gly Pro Val Leu Ser Leu Leu Thr Cys Phe His Phe Gly Arg | | | 672 |
| 210 | 215 | 220 | |
| cac cac gaa cac cat ctg agc ccc tgg cgg ccc tgg tgg cgt ctg tgg His His Glu His His Leu Ser Pro Trp Arg Pro Trp Trp Arg Leu Trp | | | 720 |
| 225 | 230 | 235 | 240 |
| cgc ggc gag tct tga Arg Gly Glu Ser | | | 735 |

<210> 36

<211> 244

<212> PRT

<213> Brevundimonas aurantiaca

<400> 36

| | | |
|---|---|----|
| Met Thr Ala Ala Val Ala Glu Pro Arg Thr Val Pro Arg Gln Thr Trp | | |
| 1 | 5 | 10 |
| | | 15 |

| | | |
|---|----|----|
| Ile Gly Leu Thr Leu Ala Gly Met Ile Val Ala Gly Trp Ala Val Leu | | |
| 20 | 25 | 30 |

His Val Tyr Gly Val Tyr Phe His Arg Trp Gly Pro Leu Thr Leu Val
35 40 45

Ile Ala Pro Ala Ile Val Ala Val Gln Thr Trp Leu Ser Val Gly Leu
50 55 60

Phe Ile Val Ala His Asp Ala Met Tyr Gly Ser Leu Ala Pro Gly Arg
65 70 75 80

Pro Arg Leu Asn Ala Ala Val Gly Arg Leu Thr Leu Gly Leu Tyr Ala
85 90 95

Gly Phe Arg Phe Asp Arg Leu Lys Thr Ala His His Ala His His Ala
100 105 110

Ala Pro Gly Thr Ala Asp Asp Pro Asp Phe His Ala Pro Ala Pro Arg
115 120 125

Ala Phe Leu Pro Trp Phe Leu Asn Phe Phe Arg Thr Tyr Phe Gly Trp
130 135 140

Arg Glu Met Ala Val Leu Thr Ala Leu Val Leu Ile Ala Leu Phe Gly
145 150 155 160

Leu Gly Ala Arg Pro Ala Asn Leu Leu Thr Phe Trp Ala Ala Pro Ala
165 170 175

Leu Leu Ser Ala Leu Gln Leu Phe Thr Phe Gly Thr Trp Leu Pro His
180 185 190

Arg His Thr Asp Gln Pro Phe Ala Asp Ala His His Ala Arg Ser Ser
195 200 205

Gly Tyr Gly Pro Val Leu Ser Leu Leu Thr Cys Phe His Phe Gly Arg
210 215 220

His His Glu His His Leu Ser Pro Trp Arg Pro Trp Trp Arg Leu Trp
225 230 235 240

Arg Gly Glu Ser

<210> 37

<211> 690

<212> DNA

<213> Nodularia spumigena NSOR10

<220>

<221> CDS

<222> (1)..(690)

<223>

<400> 37

```

atg gcg atc gcc att att agt ata tgg gct atc agc cta ggt ttg tta
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu
1           5                  10          11          15

```

48

ctt tat att gat ata tcc caa ttc aag ttt tgg atg ttg tta ccg ctc
 Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu
 20 25 30 .

96

ata ttt tgg caa aca ttt tta tat acg gga tta ttt att aca gct cat
Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
35 40 45

144

gat gcc atg cat ggg gta gtt ttt ccc aaa aat ccc aaa atc aac cat
Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His
50 55 60

192

ttc att ggc tca ttg tgc ctg ttt ctt tat ggt ctt tta cct tat caa
 Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu Leu Pro Tyr Gln
 65 70 75 80

240

aaa ctt tta aaa aag cat tgg cta cat cac cat aat cca gcc agt gaa
Lys Leu Leu Lys Lys His Trp Leu His His His Asn Pro Ala Ser Glu
85 .90 .95

288

```

aca gat cca gat ttt cac aac ggg aag cag aaa aac ttt ttt gct tgg
Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn Phe Phe Ala Trp
          100      105      110

```

336

tat tta tat ttt atg aag cgt tac tgg agt tgg tta caa att atc aca
 Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu Gln Ile Ile Thr
 115 120 125

384

```

tta atg att att tat aac tta cta aaa tat ata tgg cat ttt cca gag
Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp His Phe Pro Glu
    130           135           140

```

```

gat aat atg act tat ttt tgg gta gtt ccc tca att tta agt tct tta      480
Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu
145           150           155           160

```

caa tta ttt tat ttt gga act ttt cta ccc cac agt gag cct gta gaa 528
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu
 165 170 . 175

ggt tat aaa gag cct cat cgt tcc caa act att agc cgt ccc att tgg 576
 Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp
 180 185 190

tgg tca ttt ata act tgt tac cat ttt ggt tat cat tac gaa cat cat 624
Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His
195 200 . 205

gaa tac ccc cat gtt cct tgg tgg caa tta cca gaa att tat aaa atg 672
 Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met
 210 215 220

tct aaa tca aat ttg tga 690
Ser Lys Ser Asn Leu
225

<210> 38

<211> 229

<212> PRT

<213> Nodularia spumigena NSOR10

<400> 38

Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu
20 25 30

Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
 35 40 45

Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His
50 55 60

Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu Leu Pro Tyr Gln
 65 70 75 80

Lys Leu Leu Lys Lys His Trp Leu His His His Asn Pro Ala Ser Glu
85 90 .95

Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn Phe Phe Ala Trp
100 105 110

Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu Gln Ile Ile Thr
115 120 125

Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp His Phe Pro Glu
130 135 140

Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu
145 150 155 160

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu
165 170 175

Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp
180 185 190

Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His
195 200 205

Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met
210 215 220

Ser Lys Ser Asn Leu
225

<210> 39

<211> 789

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> CDS

<222> (1)..(789)

<223>

<400> 39

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48
 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144
 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192
 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat 240
 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
 65 70 75 80

ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca 288
 Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
 85 90 95

cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag 336
 Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
 100 105 110

aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat 384
 Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
 115 120 125

ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc 432
 Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
 130 135 140

atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta 480
 Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
 145 150 155 160

ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc 528
 Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
 165 170 175

tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat 576
 Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
 180 185 190

ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat 624

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
195 200 205

ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc 672
Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
210 215 220

gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat 720
Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768
Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

aat tca gta acc aat tcg taa 789
Asn Ser Val Thr Asn Ser
260

<210> 40

<211> 262

<212> PRT

<213> Nostoc punctiforme ATCC 29133

<400> 40

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
85 90 95

62

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

Asn Ser Val Thr Asn Ser
260

<210> 41

<211> 762

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> CDS

<222> (1)...(762)

<223>

| | | | | |
|---|-----|-----|-----|-----|
| <400> 41 | | | | |
| gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca | | | | 48 |
| Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro | | | | |
| 1 | 5 | 10 | 15 | |
| gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc | | | | 96 |
| Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val | | | | |
| 20 | 25 | 30 | | |
| att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac | | | | 144 |
| Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp | | | | |
| 35 | 40 | 45 | | |
| atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa | | | | 192 |
| Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln | | | | |
| 50 | 55 | 60 | | |
| aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat | | | | 240 |
| Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His | | | | |
| 65 | 70 | 75 | 80 | |
| ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca | | | | 288 |
| Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr | | | | |
| 85 | 90 | 95 | | |
| ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa | | | | 336 |
| Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys | | | | |
| 100 | 105 | 110 | | |
| aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat | | | | 384 |
| Lys His Trp Leu His His Asn Pro Ala Ser Ser Ile Asp Pro Asp | | | | |
| 115 | 120 | 125 | | |
| ttt cac aat ggt aaa cac caa agt ttc ttt gct ttg tat ttt cat ttt | | | | 432 |
| Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe | | | | |
| 130 | 135 | 140 | | |
| atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att | | | | 480 |
| Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile | | | | |
| 145 | 150 | 155 | 160 | |
| tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act | | | | 528 |
| Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr | | | | |
| 165 | 170 | 175 | | |
| tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat | | | | 576 |
| Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr | | | | |
| 180 | 185 | 190 | | |

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624
 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672
 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720
 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag // 762
 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys //
 245 250

<210> 42

<211> 253

<212> PRT

<213> Nostoc punctiforme ATCC 29133

<400> 42

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
 1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
 20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp
 35 40 // 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
 50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
 65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
 85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
 100 105 110

Lys His Trp Leu His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
245 250

<210> 43

<211> 1536

<212> DNA

<213> Deinococcus radiodurans R1

<220>

<221> CDS

<222> (1)..(1536)

<223>

| | | | | | | | | | | | | | | | | | |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|-----|
| <400> | 43 | | | | | | | | | | | | | | | | |
| atg | ccg | gat | tac | gac | ctg | atc | gtc | atg | ggc | gcf | ggc | cac | aac | gcf | ctg | | 48 |
| Met | Pro | Asp | Tyr | Asp | Leu | Ile | Val | Met | Gly | Ala | Gly | His | Asn | Ala | Leu | | |
| 1 | | | | 5 | | | | 10 | | | | | | | 15 | | |
| gtg | act | gct | gcc | tac | gcc | gcc | cgf | gcf | ggc | ctg | aaa | gtc | ggc | gtg | ttc | | 96 |
| Val | Thr | Ala | Ala | Tyr | Ala | Ala | Arg | Ala | Gly | Leu | Lys | Val | Gly | Val | Phe | | |
| | | | | 20 | | | | 25 | | | 30 | | | | | | |
| gag | cgf | cgf | cac | ctc | gtc | ggc | ggg | gcf | gtc | agc | acc | gag | gag | gtc | gtg | | 144 |
| Glu | Arg | Arg | His | Leu | Val | Gly | Gly | Ala | Val | Ser | Thr | Glu | Glu | Val | Val | | |
| | | | | 35 | | | | 40 | | | 45 | | | | | | |
| ccc | ggt | tac | cgc | ttc | gac | tac | ggc | ggc | agc | gcc | cac | atc | ctg | att | cgg | | 192 |
| Pro | Gly | Tyr | Arg | Phe | Asp | Tyr | Gly | Gly | Ser | Ala | His | Ile | Leu | Ile | Arg | | |
| | | | | 50 | | | | 55 | | | 60 | | | | | | |
| atg | acg | ccc | atc | gtg | cgc | gaa | ctc | gaa | ctc | acg | cgf | cac | ggg | ctg | cat | | 240 |
| Met | Thr | Pro | Ile | Val | Arg | Glu | Leu | Glu | Leu | Thr | Arg | His | Gly | Leu | His | | |
| | | | | 65 | | | | 70 | | | 75 | | | 80 | | | |
| tac | ctc | gaa | gtg | gac | cct | atg | ttt | cac | gct | tcc | gac | ggt | gaa | acg | ccc | | 288 |
| Tyr | Leu | Glu | Val | Asp | Pro | Met | Phe | His | Ala | Ser | Asp | Gly | Glu | Thr | Pro | | |
| | | | | 85 | | | | 90 | | | 95 | | | | | | |
| tgg | ttc | att | cac | cgc | gac | gcc | ggg | cgf | acc | atc | cgc | gaa | ctg | gac | gaa | | 336 |
| Trp | Phe | Ile | His | Arg | Asp | Ala | Gly | Arg | Thr | Ile | Arg | Glu | Leu | Asp | Glu | | |
| | | | | 100 | | | | 105 | | | 110 | | | | | | |
| aag | ttt | ccc | ggg | cag | ggc | gac | gcc | tac | ggg | cgc | ttt | ctc | gac | gat | tgg | | 384 |
| Lys | Phe | Pro | Gly | Gln | Gly | Asp | Ala | Tyr | Gly | Arg | Phe | Leu | Asp | Asp | Trp | | |
| | | | | 115 | | | | 120 | | | 125 | | | | | | |
| aca | ccc | ttc | gcf | cgc | gcc | gtg | gcc | gac | ctg | ttc | aat | tcg | gcf | ccg | ggg | | 432 |
| Thr | Pro | Phe | Ala | Arg | Ala | Val | Ala | Asp | Leu | Phe | Asn | Ser | Ala | Pro | Gly | | |
| | | | | 130 | | | | 135 | | | 140 | | | | | | |
| ccg | ctc | gac | ctg | ggc | aaa | atg | gtg | atg | cgc | agc | ggc | cag | ggc | aag | gac | | 480 |
| Pro | Leu | Asp | Leu | Gly | Lys | Met | Val | Met | Arg | Ser | Gly | Gln | Gly | Lys | Asp | | |
| | | | | 145 | | | | 150 | | | 155 | | | 160 | | | |
| tgg | aac | gag | cag | ctc | ccg | cgc | atc | ctg | cgf | ccc | tac | ggc | gac | gtg | gcf | | 528 |
| Trp | Asn | Glu | Gln | Leu | Pro | Arg | Ile | Leu | Arg | Pro | Tyr | Gly | Asp | Val | Ala | | |
| | | | | 165 | | | | 170 | | | 175 | | | | | | |
| cgc | gag | tac | ttc | agc | gag | gag | cgf | gtg | cgf | gct | ccc | ctg | acc | tgg | atg | | 576 |
| Arg | Glu | Tyr | Phe | Ser | Glu | Glu | Arg | Val | Arg | Ala | Pro | Leu | Thr | Trp | Met | | |
| | | | | 180 | | | | 185 | | | 190 | | | | | | |
| gcf | gcc | cag | agc | ggc | ccc | cca | ccc | tcg | gac | ccg | ctg | agc | gcf | ccc | ttt | | 624 |
| Ala | Ala | Gln | Ser | Gly | Pro | Pro | Pro | Ser | Asp | Pro | Leu | Ser | Ala | Pro | Phe | | |
| | | | | 195 | | | | 200 | | | 205 | | | | | | |
| ttg | ctg | tgg | cac | ccg | ctc | tac | cac | gaa | ggc | ggc | gtg | gcf | ccg | ccc | aaa | | 672 |

| | | | |
|---|-----|-----|------|
| Leu Leu Trp His Pro Leu Tyr His Glu Gly Gly Val Ala Arg Pro Lys | | | |
| 210 | 215 | 220 | |
| ggc ggc agc ggc ggc ctg acc aaa gcc ctg cgc cgg gcc acc gag gcc | | | 720 |
| Gly Gly Ser Gly Gly Leu Thr Lys Ala Leu Arg Arg Ala Thr Glu Ala | | | |
| 225 | 230 | 235 | 240 |
| gaa ggc ggc gag gtc ttc acc gac gcg ccg gtc aag gaa att ctg gtc | | | 768 |
| Glu Gly Gly Glu Val Phe Thr Asp Ala Pro Val Lys Glu Ile Leu Val | | | |
| 245 | 250 | 255 | |
| aag gac ggc aag gcg cag ggc atc cgg ctg gaa agc ggc gag acg tac | | | 816 |
| Lys Asp Gly Lys Ala Gln Gly Ile Arg Leu Glu Ser Gly Glu Thr Tyr | | | |
| 260 | 265 | 270 | |
| acc gcc cgc gcc gtc gtg tcg ggc gtc cac atc ctg acc act gcg aat | | | 864 |
| Thr Ala Arg Ala Val Val Ser Gly Val His Ile Leu Thr Thr Ala Asn | | | |
| 275 | 280 | 285 | |
| gcc ctg ccc gcc gaa tat gtc cct agc gcc agg aat gtg cgc gtg | | | 912 |
| Ala Leu Pro Ala Glu Tyr Val Pro Ser Ala Ala Arg Asn Val Arg Val | | | |
| 290 | 295 | 300 | |
| ggc aac ggc ttc ggc atg att ttg cgc ctc gcc ctc agt gaa aaa gtc | | | 960 |
| Gly Asn Gly Phe Gly Met Ile Leu Arg Leu Ala Leu Ser Glu Lys Val | | | |
| 305 | 310 | 315 | 320 |
| aaa tac cgt cac cac acc gag ccc gac tca cgc atc ggc ctg gga ttg | | | 1008 |
| Lys Tyr Arg His His Thr Glu Pro Asp Ser Arg Ile Gly Leu Gly Leu | | | |
| 325 | 330 | 335 | |
| ctg atc aaa aac gag cgg caa atc atg cag ggc tac ggc gaa tac ctc | | | 1056 |
| Leu Ile Lys Asn Glu Arg Gln Ile Met Gln Gly Tyr Gly Glu Tyr Leu | | | |
| 340 | 345 | 350 | |
| gcc ggg cag ccc acc acc gac ccg ccc ctc gtc gcc atg agc ttc agc | | | 1104 |
| Ala Gly Gln Pro Thr Thr Asp Pro Pro Leu Val Ala Met Ser Phe Ser | | | |
| 355 | 360 | 365 | |
| gcg gtg gac gac tcg ctc gcc cca ccg aac ggc gac gtg ttg tgg ctg | | | 1152 |
| Ala Val Asp Asp Ser Leu Ala Pro Pro Asn Gly Asp Val Leu Trp Leu | | | |
| 370 | 375 | 380 | |
| tgg gcg cag tac tac ccc ttc gag ctc gcc acc ggg agc tgg gaa acg | | | 1200 |
| Trp Ala Gln Tyr Tyr Pro Phe Glu Leu Ala Thr Gly Ser Trp Glu Thr | | | |
| 385 | 390 | 395 | 400 |
| cgc acc gcc gaa gcg cgg gag aac atc ctg cgg gcc ttt gag cac tac | | | 1248 |
| Arg Thr Ala Glu Ala Arg Glu Asn Ile Leu Arg Ala Phe Glu His Tyr | | | |
| 405 | 410 | 415 | |
| gcg ccg ggc acc cgc gac acg att gtg ggc gaa ctc gtg cag acg ccg | | | 1296 |
| Ala Pro Gly Thr Arg Asp Thr Ile Val Gly Glu Leu Val Gln Thr Pro | | | |
| 420 | 425 | 430 | |
| cag tgg ctg gaa acc aac ctc ggc ctg cac cgg ggc aac gtg atg cac | | | 1344 |

Gln Trp Leu Glu Thr Asn Leu Gly Leu His Arg Gly Asn Val Met His
 435 440 445

ctg gaa atg tcc ttc gac cag atg ttc tcc ttc cgc ccc tgg ctg aaa 1392
 Leu Glu Met Ser Phe Asp Gln Met Phe Ser Phe Arg Pro Trp Leu Lys
 450 455 460

gcg agc cag tac cgc tgg ccg ggc gtg cag ggg ctg tac ctc acc ggc 1440
 Ala Ser Gln Tyr Arg Trp Pro Gly Val Gln Gly Leu Tyr Leu Thr Gly
 465 470 475 480

gcc agc acc cac ccc ggc gga ggc atc atg ggc gcc tcg gga cgc aac 1488
 Ala Ser Thr His Pro Gly Gly Ile Met Gly Ala Ser Gly Arg Asn
 485 490 495

gcg gcg cgg gtc atc gtg aag gac ctg acg cgg agg cgc tgg aaa tga 1536
 Ala Ala Arg Val Ile Val Lys Asp Leu Thr Arg Arg Arg Trp Lys
 500 505 510

<210> 44

<211> 511

<212> PRT

<213> Deinococcus radiodurans R1

<400> 44

Met Pro Asp Tyr Asp Leu Ile Val Met Gly Ala Gly His Asn Ala Leu
 1 5 10 15

Val Thr Ala Ala Tyr Ala Ala Arg Ala Gly Leu Lys Val Gly Val Phe
 20 25 30

Glu Arg Arg His Leu Val Gly Gly Ala Val Ser Thr Glu Glu Val Val
 35 40 45

Pro Gly Tyr Arg Phe Asp Tyr Gly Gly Ser Ala His Ile Leu Ile Arg
 50 55 60

Met Thr Pro Ile Val Arg Glu Leu Glu Leu Thr Arg His Gly Leu His
 65 70 75 80

Tyr Leu Glu Val Asp Pro Met Phe His Ala Ser Asp Gly Glu Thr Pro
 85 90 95

Trp Phe Ile His Arg Asp Ala Gly Arg Thr Ile Arg Glu Leu Asp Glu
100 105 110

Lys Phe Pro Gly Gln Gly Asp Ala Tyr Gly Arg Phe Leu Asp Asp Trp
115 120 125

Thr Pro Phe Ala Arg Ala Val Ala Asp Leu Phe Asn Ser Ala Pro Gly
130 135 140

Pro Leu Asp Leu Gly Lys Met Val Met Arg Ser Gly Gln Gly Lys Asp
145 150 155 160

Trp Asn Glu Gln Leu Pro Arg Ile Leu Arg Pro Tyr Gly Asp Val Ala
165 170 175

Arg Glu Tyr Phe Ser Glu Glu Arg Val Arg Ala Pro Leu Thr Trp Met
180 185 190

Ala Ala Gln Ser Gly Pro Pro Pro Ser Asp Pro Leu Ser Ala Pro Phe
195 200 205

Leu Leu Trp His Pro Leu Tyr His Glu Gly Gly Val Ala Arg Pro Lys
210 215 220

Gly Gly Ser Gly Gly Leu Thr Lys Ala Leu Arg Arg Ala Thr Glu Ala
225 230 235 240

Glu Gly Gly Glu Val Phe Thr Asp Ala Pro Val Lys Glu Ile Leu Val
245 250 255

Lys Asp Gly Lys Ala Gln Gly Ile Arg Leu Glu Ser Gly Glu Thr Tyr
260 265 270

Thr Ala Arg Ala Val Val Ser Gly Val His Ile Leu Thr Thr Ala Asn
275 280 285

Ala Leu Pro Ala Glu Tyr Val Pro Ser Ala Ala Arg Asn Val Arg Val
290 295 300

Gly Asn Gly Phe Gly Met Ile Leu Arg Leu Ala Leu Ser Glu Lys Val
305 310 315 320

70

Lys Tyr Arg His His Thr Glu Pro Asp Ser Arg Ile Gly Leu Gly Leu
325 330 335

Leu Ile Lys Asn Glu Arg Gln Ile Met Gln Gly Tyr Gly Glu Tyr Leu
340 345 350

Ala Gly Gln Pro Thr Thr Asp Pro Pro Leu Val Ala Met Ser Phe Ser
355 360 365

Ala Val Asp Asp Ser Leu Ala Pro Pro Asn Gly Asp Val Leu Trp Leu
370 375 380

Trp Ala Gln Tyr Tyr Pro Phe Glu Leu Ala Thr Gly Ser Trp Glu Thr
385 390 395 400

Arg Thr Ala Glu Ala Arg Glu Asn Ile Leu Arg Ala Phe Glu His Tyr
405 410 415

Ala Pro Gly Thr Arg Asp Thr Ile Val Gly Glu Leu Val Gln Thr Pro
420 425 430

Gln Trp Leu Glu Thr Asn Leu Gly Leu His Arg Gly Asn Val Met His
435 440 445

Leu Glu Met Ser Phe Asp Gln Met Phe Ser Phe Arg Pro Trp Leu Lys
450 455 460

Ala Ser Gln Tyr Arg Trp Pro Gly Val Gln Gly Leu Tyr Leu Thr Gly
465 470 475 480

Ala Ser Thr His Pro Gly Gly Ile Met Gly Ala Ser Gly Arg Asn
485 490 495

Ala Ala Arg Val Ile Val Lys Asp Leu Thr Arg Arg Arg Trp Lys
500 505 510

<210> 45

<211> 1650

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (112) . . (1614)

<223>

<400> 45

ggcacgagga aacttttctc tcttcactag ctgtttacat gcttcaaatt tcaagatttt 60

aggaccccat ttgaagttt cttgaaacaa atattaccct gttggaaaaa g atg gat 117
Met Asp
1

act ttg ttg aaa acc cca aat aac ctt gaa ttt ctg aac cca cat cat 1655
Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro His His
5 10 15

ggt ttt gct gtt aaa gct agt acc ttt aga tct gag aag cat cat aat 213
 Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His His Asn
 20 25 30

ttt ggt tct agg aag ttt tgt gaa act ttg ggt aga agt gtt tgt gtt 261
 Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val Cys Val
 35 40 45 50

gag aat ctt gat ttt gag ctt cct atg tat gac cct tca aaa ggg gtt 357
 Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val
 70 75 80

gtt gtg gat ctt gct gtg gtt ggt ggc cct gca gga ctt gct gtt 405
 Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu Ala Val
 85 90 95

gca cag caa gtt tct gaa gca gga ctc tct gtt tgt tca att gat ccg 453
 Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro
 100 105 110

aat cct aaa ttg ata tgg cct aat aac tat ggt gtt tgg gtg gat gaa 501
Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu
115 120 125 130

ttt gag gct atg gac ttg tta gat tgt cta gat gct acc tgg tct ggt 549
Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly
135 140 145

gca gca gtq tac att qat qat aat acq qct aaa qat ctt cat aqa cct 597

| | | | |
|---|-----|-----|------|
| Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His Arg Pro | | | |
| 150 | 155 | 160 | |
| tat gga agg gtt aac cgg aaa cag ctg aaa tcg aaa atg atg cag aaa | | | 645 |
| Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys | | | |
| 165 | 170 | 175 | |
| tgt ata atg aat ggt gtt aaa ttc cac caa gcc aaa gtt ata aag gtg | | | 693 |
| Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile Lys Val | | | |
| 180 | 185 | 190 | |
| att cat gag gaa tcg aaa tcc atg ttg ata tgc aat gat ggt att act | | | 741 |
| Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly Ile Thr | | | |
| 195 | 200 | 205 | 210 |
| att cag gca acg gtg gtg ctc gat gca act ggc ttc tct aga tct ctt | | | 789 |
| Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu | | | |
| 215 | 220 | 225 | |
| gtt cag tat gat aag cct tat aac ccc ggg tat caa gtt gct tat ggc | | | 837 |
| Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly | | | |
| 230 | 235 | 240 | |
| att ttg gct gaa gtg gaa gag cac ccc ttt gat gta aac aag atg gtt | | | 885 |
| Ile Leu Ala Glu Val Glu His Pro Phe Asp Val Asn Lys Met Val | | | |
| 245 | 250 | 255 | |
| ttc atg gat tgg cga gat tct cat ttg aag aac aat act gat ctc aag | | | 933 |
| Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp Leu Lys | | | |
| 260 | 265 | 270 | |
| gag aga aat agt aga ata cca act ttt ctt tat gca atg cca ttt tca | | | 981 |
| Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser | | | |
| 275 | 280 | 285 | 290 |
| tcc aac agg ata ttt ctt gaa gaa aca tca ctc gta gct cgt cct ggc | | | 1029 |
| Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly | | | |
| 295 | 300 | 305 | |
| ttg cgt ata gat gat att caa gaa cga atg gtg gct cgt tta aac cat | | | 1077 |
| Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu Asn His | | | |
| 310 | 315 | 320 | |
| ttg ggg ata aaa gtg aag agc att gaa gaa gat gaa cat tgt cta ata | | | 1125 |
| Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Leu Ile | | | |
| 325 | 330 | 335 | |
| cca atg ggt ggt cca ctt cca gta tta cct cag aga gtc gtt gga atc | | | 1173 |
| Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile | | | |
| 340 | 345 | 350 | |
| ggt ggt aca gct ggc atg gtt cat cca tcc acc ggt tat atg gtg gca | | | 1221 |
| Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala | | | |
| 355 | 360 | 365 | 370 |
| agg aca cta gct gcg gct cct gtt gtc aat gcc ata att caa tac | | | 1269 |

73

Arg Thr Leu Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr
 375 380 385
 ctc ggt tct gaa aga agt cat tcg ggt aat gaa tta tcc aca gct gtt 1317
 Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr Ala Val
 390 395 400
 tgg aaa gat ttg tgg cct ata gag agg aga cgt caa aga gag ttc ttc 1365
 Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe
 405 410 415
 tgc ttc ggt atg gat att ctt ctg aag ctt gat tta cct gct aca aga 1413
 Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg
 420 425 430
 agg ttc ttt gat gca ttc ttt gac tta gaa cct cgt tat tgg cat ggc 1461
 Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly
 435 440 445 450
 ttc tta tcg tct cga ttg ttt cta cct gaa ctc ata gtt ttt ggg ctg 1509
 Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu
 455 460 465
 tct cta ttc tct cat gct tca aat act tct aga ttt gag ata atg aca 1557
 Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile Met Thr
 470 475 480
 aag gga act gtt cca tta gta aat atg atc aac aat ttg tta cag gat 1605
 Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu Gln Asp
 485 490 495
 aaa gaa tga atcccgagtaa ttccggatatct tgtccaatct cgtgcc 1650
 Lys Glu
 500
 <210> 46
 <211> 500
 <212> PRT
 <213> Lycopersicon esculentum
 <400> 46
 Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro
 1 5 10 15
 His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His
 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val
35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr
50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys
65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu
85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile
100 105 110

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val
115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp
130 135 140

Ser Gly Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His
145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met.
165 170 175

Gln Lys Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile
180 185 190

Lys Val Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly
195 200 205

Ile Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg
210 215 220

Ser Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala
225 230 235 240

Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys
245 250 255

75

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp
260 265 270

Leu Lys Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro
275 280 285

Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg
290 295 300

Pro Gly Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu
305 310 315 320

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys
325 330 335

Leu Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val
340 345 350

Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met
355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile
370 375 380

Gln Tyr Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr
385 390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu
405 410 415

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala
420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp
435 440 445

His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe
450 455 460

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile
465 470 475 480

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu
485 490 495

Gln Asp Lys Glu
500

<210> 47

<211> 1666

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> CDS

<222> (1)..(1494)

<223>

<400> 47

atg gaa gct ctt ctc aag cct ttt cca tct ctt tta ctt tcc tct cct
Met Glu Ala Leu Leu Lys Pro Phe Pro Ser Leu Leu Leu Ser Ser Pro
1 5 10 15

48

```

aca ccc cat agg tct att ttc caa caa aat ccc tct ttt cta agt ccc
Thr Pro His Arg Ser Ile Phe Gln Gln Asn Pro Ser Phe Leu Ser Pro
20          25          30

```

96

acc acc aaa aaa aaa tca aga aaa tgt ctt ctt aga aac aaa agt agt
 Thr Thr Lys Lys Lys Ser Arg Lys Cys Leu Leu Arg Asn Lys Ser Ser
 35 40 45

144

```

aaa ctt ttt tgt agc ttt ctt gat tta gca ccc aca tca aag cca gag
Lys Leu Phe Cys Ser Phe Leu Asp Leu Ala Pro Thr Ser Lys Pro Glu
      50           55           60

```

192

tct tta gat gtt aac atc tca tgg gtt gat cct aat tcg aat cggt
Ser Leu Asp Val Asn Ile Ser Trp Val Asp Pro Asn Ser Asn Arg Ala
65 70 75 80

240

caa ttc gac gtg atc att atc gga gct ggc cct gct ggg ctc agg cta
 Gln Phe Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Arg Leu
 85 90 95

288

```

gct gaa caa gtt tct aaa tat ggt att aag gta tgt tgt gtt gac cct
Ala Glu Gln Val Ser Lys Tyr Gly Ile Lys Val Cys Cys Val Asp Pro
          100           105           110

```

| | |
|---|------|
| tca cca ctc tcc atg tgg cca aat aat tat ggt gtt tgg gtt gat gag Ser Pro Leu Ser Met Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu 115 120 125 | 384 |
| ttt gag aat tta gga ctg gaa aat tgt tta gat cat aaa tgg cct atg Phe Glu Asn Leu Gly Leu Glu Asn Cys Leu Asp His Lys Trp Pro Met 130 135 140 | 432 |
| act tgt gtg cat ata aat gat aac aaa act aag tat ttg gga aga cca Thr Cys Val His Ile Asn Asp Asn Lys Thr Lys Tyr Leu Gly Arg Pro 145 150 155 160 | 480 |
| tat ggt aga gtt agt aga aag aag ctg aag ttg aaa ttg ttg aat agt Tyr Gly Arg Val Ser Arg Lys Lys Leu Lys Leu Lys Leu Asn Ser 165 170 175 | 528 |
| tgt gtt gag aac aga gtg aag ttt tat aaa gct aag gtt tgg aaa gtg Cys Val Glu Asn Arg Val Lys Phe Tyr Lys Ala Lys Val Trp Lys Val 180 185 190 | 576 |
| gaa cat gaa gaa ttt gag tct tca att gtt tgt gat gat ggt aag aag Glu His Glu Glu Phe Glu Ser Ser Ile Val Cys Asp Asp Gly Lys Lys 195 200 205 | 624 |
| ata aga ggt agt ttg gtt gtg gat gca agt ggt ttt gct agt gat ttt Ile Arg Gly Ser Leu Val Val Asp Ala Ser Gly Phe Ala Ser Asp Phe 210 215 220 | 672 |
| ata gag tat gac agg cca aga aac cat ggt tat caa att gct cat ggg Ile Glu Tyr Asp Arg Pro Arg Asn His Gly Tyr Gln Ile Ala His Gly 225 230 235 240 | 720 |
| gtt tta gta gaa gtt gat aat cat cca ttt gat ttg gat aaa atg gtg Val Leu Val Glu Val Asp Asn His Pro Phe Asp Leu Asp Lys Met Val 245 250 255 | 768 |
| ctt atg gat tgg agg gat tct cat ttg ggt aat gag cca tat tta agg Leu Met Asp Trp Arg Asp Ser His Leu Gly Asn Glu Pro Tyr Leu Arg 260 265 270 | 816 |
| gtg aat aat gct aaa gaa cca aca ttc ttg tat gca atg cca ttt gat Val Asn Asn Ala Lys Glu Pro Thr Phe Leu Tyr Ala Met Pro Phe Asp 275 280 285 | 864 |
| aga gat ttg gtt ttc ttg gaa gag act tct ttg gtg agt cgt cct gtt Arg Asp Leu Val Phe Leu Glu Glu Thr Ser Leu Val Ser Arg Pro Val 290 295 300 | 912 |
| tta tcg tat atg gaa gta aaa aga agg atg gtg gca aga tta agg cat Leu Ser Tyr Met Glu Val Lys Arg Arg Met Val Ala Arg Leu Arg His 305 310 315 320 | 960 |
| ttg ggg atc aaa gtg aaa agt gtt att gag gaa gag aaa tgt gtg atc Leu Gly Ile Lys Val Lys Ser Val Ile Glu Glu Glu Lys Cys Val Ile 325 330 335 | 1008 |

| | |
|---|-----------|
| cct atg gga gga cca ctt ccg cg att cct caa aat gtt atg gct att Pro Met Gly Gly Pro Leu Pro Arg Ile Pro Gln Asn Val Met Ala Ile 340 345 350 | 1056 |
| ggt ggg aat tca ggg ata gtt cat cca tca aca ggg tac atg gtg gct Gly Gly Asn Ser Gly Ile Val His Pro Ser Thr Gly Tyr Met Val Ala 355 360 365 | 1104 |
| agg agc atg gct tta gca cca gta cta gct gaa gcc atc gtc gag ggg Arg Ser Met Ala Leu Ala Pro Val Leu Ala Glu Ala Ile Val Glu Gly 370 375 380 | 1152 |
| ctt ggc tca aca aga atg ata aga ggg tct caa ctt tac cat aga gtt Leu Gly Ser Thr Arg Met Ile Arg Gly Ser Gln Leu Tyr His Arg Val 385 390 395 400 | 1200 |
| tgg aat ggt ttg tgg cct ttg gat aga aga tgt gtt aga gaa tgt tat Trp Asn Gly Leu Trp Pro Leu Asp Arg Arg Cys Val Arg Glu Cys Tyr 405 410 415 | 1248 |
| tca ttt ggg atg gag aca ttg ttg aag ctt gat ttg aaa ggg act agg Ser Phe Gly Met Glu Thr Leu Leu Lys Leu Asp Leu Lys Gly Thr Arg 420 425 430 | 1296 |
| aga ttg ttt gac gct ttc ttt gat ctt gat cct aaa tac tgg caa ggg Arg Leu Phe Asp Ala Phe Phe Asp Leu Asp Pro Lys Tyr Trp Gln Gly 435 440 445 | 1344 |
| ttc ctt tct tca aga ttg tct gtc aaa gaa ctt ggt tta ctc agc ttg Phe Leu Ser Ser Arg Leu Ser Val Lys Glu Leu Gly Leu Leu Ser Leu 450 455 460 | 1392 |
| tgt ctt ttc gga cat ggc tca aac atg act agg ttg gat att gtt aca Cys Leu Phe Gly His Gly Ser Asn Met Thr Arg Leu Asp Ile Val Thr 465 470 475 480 | 1440 |
| aaa tgt cct ctt cct ttg gtt aga ctg att ggc aat cta gca ata gag Lys Cys Pro Leu Pro Leu Val Arg Leu Ile Gly Asn Leu Ala Ile Glu 485 490 495 | 1488 |
| agc ctt tgaatgtgaa aagtttgaat cattttcttc attttaattt ctggattat Ser Leu | 1544 |
| tttcatattt tctcaattgc aaaagtgaga taagagctac atactgtcaa caaataaact actattggaa agttaaaata tgtgtttgtt gtatgttatt ctaatggaat ggattttgta | 1604 1664 |
| aa | 1666 |
| <210> 48 | |
| <211> 498 | |

<212> PRT

<213> Lycopersicon esculentum

<400> 48

Met Glu Ala Leu Leu Lys Pro Phe Pro Ser Leu Leu Leu Ser Ser Pro
1 5 10 15

Thr Pro His Arg Ser Ile Phe Gln Gln Asn Pro Ser Phe Leu Ser Pro
20 25 30

Thr Thr Lys Lys Lys Ser Arg Lys Cys Leu Leu Arg Asn Lys Ser Ser
35 40 45

Lys Leu Phe Cys Ser Phe Leu Asp Leu Ala Pro Thr Ser Lys Pro Glu
50 55 60

Ser Leu Asp Val Asn Ile Ser Trp Val Asp Pro Asn Ser Asn Arg Ala
65 70 75 80

Gln Phe Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Arg Leu
85 90 95

Ala Glu Gln Val Ser Lys Tyr Gly Ile Lys Val Cys Cys Val Asp Pro
100 105 110

Ser Pro Leu Ser Met Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu
115 120 125

Phe Glu Asn Leu Gly Leu Glu Asn Cys Leu Asp His Lys Trp Pro Met
130 135 140

Thr Cys Val His Ile Asn Asp Asn Lys Thr Lys Tyr Leu Gly Arg Pro
145 150 155 160

Tyr Gly Arg Val Ser Arg Lys Lys Leu Lys Leu Lys Leu Asn Ser
165 170 175

Cys Val Glu Asn Arg Val Lys Phe Tyr Lys Ala Lys Val Trp Lys Val
180 185 190

Glu His Glu Glu Phe Glu Ser Ser Ile Val Cys Asp Asp Gly Lys Lys
195 200 205

Ile Arg Gly Ser Leu Val Val Asp Ala Ser Gly Phe Ala Ser Asp Phe
210 215 220

Ile Glu Tyr Asp Arg Pro Arg Asn His Gly Tyr Gln Ile Ala His Gly
225 230 235 240

Val Leu Val Glu Val Asp Asn His Pro Phe Asp Leu Asp Lys Met Val
245 250 255

Leu Met Asp Trp Arg Asp Ser His Leu Gly Asn Glu Pro Tyr Leu Arg
260 265 270

Val Asn Asn Ala Lys Glu Pro Thr Phe Leu Tyr Ala Met Pro Phe Asp
275 280 285

Arg Asp Leu Val Phe Leu Glu Glu Thr Ser Leu Val Ser Arg Pro Val
290 295 300

Leu Ser Tyr Met Glu Val Lys Arg Arg Met Val Ala Arg Leu Arg His
305 310 315 320

Leu Gly Ile Lys Val Lys Ser Val Ile Glu Glu Glu Lys Cys Val Ile
325 330 335

Pro Met Gly Gly Pro Leu Pro Arg Ile Pro Gln Asn Val Met Ala Ile
340 345 350

Gly Gly Asn Ser Gly Ile Val His Pro Ser Thr Gly Tyr Met Val Ala
355 360 365

Arg Ser Met Ala Leu Ala Pro Val Leu Ala Glu Ala Ile Val Glu Gly
370 375 380

Leu Gly Ser Thr Arg Met Ile Arg Gly Ser Gln Leu Tyr His Arg Val
385 390 395 400

Trp Asn Gly Leu Trp Pro Leu Asp Arg Arg Cys Val Arg Glu Cys Tyr
405 410 415

Ser Phe Gly Met Glu Thr Leu Leu Lys Leu Asp Leu Lys Gly Thr Arg
 420 425 430

Arg Leu Phe Asp Ala Phe Phe Asp Leu Asp Pro Lys Tyr Trp Gln Gly
 435 440 445

Phe Leu Ser Ser Arg Leu Ser Val Lys Glu Leu Gly Leu Leu Ser Leu
 450 455 460

Cys Leu Phe Gly His Gly Ser Asn Met Thr Arg Leu Asp Ile Val Thr
 465 470 475 480

Lys Cys Pro Leu Pro Leu Val Arg Leu Ile Gly Asn Leu Ala Ile Glu
 485 490 495

Ser Leu

<210> 49

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3)..(971)

<223>

<400> 49

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47
 Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile
 1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95
 Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu
 20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143
 Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala
 35 40 45

| | |
|---|-----|
| aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt | 863 |
| Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly | |
| 275 280 285 | |

```

ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att      911
Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile
290           295           300

```

cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg 959
 Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp
 305 310 315

tcc aag cgg tag ggtgcggAAC caggcacGCT ggTTTcacAC ctcatgcCTG 1011
Ser Lys Arg
320

tgataagggtg tggcttagagc gatgcgtgtg agacgggtat gtcacggtcg actgggtctga 1071

tggccaatgg catcgccat gtctggcat cacgggctgg ttgcctgggt gaaggtgatg 1131

cacatcatca tgtgcgggtt gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191

caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgccc 1251

Statistical techniques for analyzing environmental data 181

catatgtac tctgtcatggat gtatttggatc gaggatggat ctggatggat ctgttattata 1431

ggctcggtgcc aqaaaatcqgtq aqtgcgtatqac tqtgcacacta tacatttcgat ccaggttgata 1551

tgcactgtct cgattgtaaa atacattcaq atgcaaaaaaaaaaaaaaaaaaaaaaa 1608

<210> 50

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 50

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser
20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val
245 250 255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys
260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly
275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro
290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser
305 310 315 320

Lys Arg

<210> 51

<211> 1163

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(942)

<223>

<400> 51

att cgg cac gag att tca gcc tcc gct agt tcc cga acc att cgc ctc 48
Ile Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu
1 5 10 15

cgt cat aac ccg ttt ctc agt cca aaa tcc gcc tca acc gcc ccg ccg 96
Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro
20 25 30

gtt ctg ttc ttc tct ccg tta act cgc aat ttt ggc gca att ttg ctg 144
Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu
35 40 45

tct aga aga aag ccg aga ttg gcg gtt tgt ttt gtg ctg gag aat gag 192

| | | |
|---|-----|-----|
| Ser Arg Arg Lys Pro Arg Leu Ala Val Cys Phe Val Leu Glu Asn Glu | | |
| 50 | 55 | 60 |
| aaa ttg aat agt act atc gaa agt gag agt gaa gta ata gag gat cgg | | 240 |
| Lys Leu Asn Ser Thr Ile Glu Ser Glu Ser Glu Val Ile Glu Asp Arg | | |
| 65 | 70 | 75 |
| ata caa gta gag att aat gag gag aag agt tta gct gcc agt tgg ctg | | 288 |
| Ile Gln Val Glu Ile Asn Glu Glu Lys Ser Leu Ala Ala Ser Trp Leu | | |
| 85 | 90 | 95 |
| gcg gag aaa ttg gcg agg aag aaa tcg gag agg ttt act tat ctt gtg | | 336 |
| Ala Glu Lys Leu Ala Arg Lys Lys Ser Glu Arg Phe Thr Tyr Leu Val | | |
| 100 | 105 | 110 |
| gca gct gtg atg tct agt ttg ggg att act tct atg gcg att ttg gcg | | 384 |
| Ala Ala Val Met Ser Ser Leu Gly Ile Thr Ser Met Ala Ile Leu Ala | | |
| 115 | 120 | 125 |
| gtt tat tac aga ttt tca tgg caa atg gag ggt gga gaa gtg cct ttt | | 432 |
| Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly Glu Val Pro Phe | | |
| 130 | 135 | 140 |
| tct gaa atg tta gct aca ttc act ctc tcg ttt ggc gct gcc gta gga | | 480 |
| Ser Glu Met Leu Ala Thr Phe Thr Leu Ser Phe Gly Ala Ala Val Gly | | |
| 145 | 150 | 155 |
| atg gag tac tgg gcg aga tgg gct cat aga gca cta tgg cat gct tct | | 528 |
| Met Glu Tyr Trp Ala Arg Trp Ala His Arg Ala Leu Trp His Ala Ser | | |
| 165 | 170 | 175 |
| tta tgg cac atg cac gag tcg cac cat aga cca aga gaa gga cct ttt | | 576 |
| Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe | | |
| 180 | 185 | 190 |
| gag atg aac gac gtt ttc gcc ata aca aat gct gtt cca gct ata ggt | | 624 |
| Glu Met Asn Asp Val Phe Ala Ile Thr Asn Ala Val Pro Ala Ile Gly | | |
| 195 | 200 | 205 |
| ctt ctt tcc tac ggt ttc ttc cat aaa ggg atc gtc cct ggc ctc tgt | | 672 |
| Leu Leu Ser Tyr Gly Phe Phe His Lys Gly Ile Val Pro Gly Leu Cys | | |
| 210 | 215 | 220 |
| ttc ggc gct gga ttg ggg atc aca gta ttt ggg atg gct tac atg ttc | | 720 |
| Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe | | |
| 225 | 230 | 235 |
| gtt cac gat gga ctg gtt cat aag aga ttt ccc gta ggg cct att gcc | | 768 |
| Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala | | |
| 245 | 250 | 255 |
| aac gtg cct tac ttt cgg agg gta gct gca gca cat cag ctt cat cac | | 816 |
| Asn Val Pro Tyr Phe Arg Arg Val Ala Ala His Gln Leu His His | | |
| 260 | 265 | 270 |
| tcg gac aaa ttt gat ggt gtc cca tat ggc ttg ttt cta gga cct aag | | 864 |

Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys
275 280 285

gaa ttg gaa gaa gta gga gga ctt gaa gag tta gaa aag gaa gtc aac 912
 Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Val Asn
 290 295 300

cga agg att aaa att tct aag gga tta tta tgatcaaaaag atacgtctga. 962
Arg Arg Ile Lys Ile Ser Lys Gly Leu Leu
305 310

taataataaa atgcgattgt atttaggctg tagattatta ttggggaaaaaa gatagaaaga 1022

tatataatatg aatataaatat aaaatgcac aagtttcta tggagaagac cttttttttt 1082

ttggtaacctg tacgtaaaaag gtgaacaatt tgatgtcccta gtactttatcg acaaaccaga 1142

agaacgataa ttcaaaaacaa a . . . 1163

<210> 52

<211> 314

<212> PRT

<213> *Lycopersicon esculentum*

<400> 52

Ile Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu
1 . . 5 10 15

Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro
20 25 30

Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu
35 40 45

Ser Arg Arg Lys Pro Arg Leu Ala Val Cys Phe Val Leu Glu Asn Glu
50 55 60 /

Lys Leu Asn Ser Thr Ile Glu Ser Glu Ser Glu Val Ile Glu Asp Arg
65 70 75 80

Ile Gln Val Glu Ile Asn Glu Glu Lys Ser Leu Ala Ala Ser Trp Leu
85 90 95

Ala Glu Lys Leu Ala Arg Lys Lys Ser Glu Arg Phe Thr Tyr Leu Val
100 105 110

Ala Ala Val Met Ser Ser Leu Gly Ile Thr Ser Met Ala Ile Leu Ala
115 120 125

Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly Glu Val Pro Phe
130 135 140

Ser Glu Met Leu Ala Thr Phe Thr Leu Ser Phe Gly Ala Ala Val Gly
145 150 155 160

Met Glu Tyr Trp Ala Arg Trp Ala His Arg Ala Leu Trp His Ala Ser
165 170 175

Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe
180 185 190

Glu Met Asn Asp Val Phe Ala Ile Thr Asn Ala Val Pro Ala Ile Gly
195 200 205

Leu Leu Ser Tyr Gly Phe Phe His Lys Gly Ile Val Pro Gly Leu Cys
210 215 220

Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe
225 230 235 240

Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala
245 250 255

Asn Val Pro Tyr Phe Arg Arg Val Ala Ala Ala His Gln Leu His His
260 265 270

Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys
275 280 285

Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Val Asn
290 295 300

Arg Arg Ile Lys Ile Ser Lys Gly Leu Leu
305 310

<210> 53

<211> 1779

<212> DNA

<213> *Arabidopsis thaliana*

<220>

<221> CDS

<222> (1)..(1779)

<223>

<400> 53

atg gat ctc cgt cgg agg cct cct aaa cca ccg gtt acc aac aac aac
 Met Asp Leu Arg Arg Arg Pro Pro Lys Pro Pro Val Thr Asn Asn Asn
 1 5 10 15

48

aac tcc aac gga tct ttc cgt tct tat cag cct cgc act tcc gat gac
 Asn Ser Asn Gly Ser Phe Arg Ser Tyr Gln Pro Arg Thr Ser Asp Asp
 20 25 30

96

gat cat cgt cgc cgg gct aca aca att gct cct cca ccg aaa gca tcc
 Asp His Arg Arg Arg Ala Thr Thr Ile Ala Pro Pro Pro Lys Ala Ser
 35 40 45

144

gac gcg ctt cct ctt ccg tta tat ctc aca aac gcc gtt ttc ttc acg
 Asp Ala Leu Pro Leu Pro Tyr Leu Thr Asn Ala Val Phe Phe Thr
 50 55 60

192

ctc ttc ttc tcc gtc gcg tat tac ctc ctc cac ccg tgg cgt gac aag
 Leu Phe Phe Ser Val Ala Tyr Tyr Leu Leu His Arg Trp Arg Asp Lys
 65 70 75 80

240

atc cgt tac aat acg cct ctt cac gtc gtc act atc aca gaa ctc ggc
 Ile Arg Tyr Asn Thr Pro Leu His Val Val Thr Ile Thr Glu Leu Gly
 85 90 95

288

gcc att att gct ctc atc gct tcg ttt atc tat ctc cta ggg ttt ttt
 Ala Ile Ile Ala Leu Ile Ala Ser Phe Ile Tyr Leu Leu Gly Phe Phe
 100 105 110

336

ggt att gac ttt gtt cag tca ttt atc tca cgt gcc tct ggt gat gct
 Gly Ile Asp Phe Val Gln Ser Phe Ile Ser Arg Ala Ser Gly Asp Ala
 115 120 125

384

tgg gat ctc gcc gat acg atc gat gat gat gac cac ccg ctt gtc acg
 Trp Asp Leu Ala Asp Thr Ile Asp Asp Asp His Arg Leu Val Thr
 130 135 140

432

| | |
|---|------|
| tgc tct cca ccg act ccg atc gtt tcc gtt gct aaa tta cct aat ccg Cys Ser Pro Pro Thr Pro Ile Val Ser Val Ala Lys Leu Pro Asn Pro 145 150 155 160 | 480 |
| gaa cct att gtt acc gaa tcg ctt cct gag gaa gac gag gag att gtg Glu Pro Ile Val Thr Glu Ser Leu Pro Glu Glu Asp Glu Glu Ile Val 165 170 175 | 528 |
| aaa tcg gtt atc gac gga gtt att cca tcg tac tcg ctt gaa tct cgt Lys Ser Val Ile Asp Gly Val Ile Pro Ser Tyr Ser Leu Glu Ser Arg 180 185 190 | 576 |
| ctc ggt gat tgc aaa aga gcg gcg tcg att cgt cgt gag gcg ttg cag Leu Gly Asp Cys Lys Arg Ala Ser Ile Arg Arg Glu Ala Leu Gln 195 200 205 | 624 |
| aga gtc acc ggg aga tcg att gaa ggg tta ccg ttg gat gga ttt gat Arg Val Thr Gly Arg Ser Ile Glu Gly Leu Pro Leu Asp Gly Phe Asp 210 215 220 | 672 |
| tat gaa tcg att ttg ggg caa tgc tgt gag atg cct gtt gga tac att Tyr Glu Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Val Gly Tyr Ile 225 230 235 240 | 720 |
| cag att cct gtt ggg att gct ggt cca ttg ttg ctt gat ggt tat gag Gln Ile Pro Val Gly Ile Ala Gly Pro Leu Leu Asp Gly Tyr Glu 245 250 255 | 768 |
| tac tct gtt cct atg gct aca acc gaa ggt tgt ttg gtt gct agc act Tyr Ser Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr 260 265 270 | 816 |
| aac aga ggc tgc aag gct atg ttt atc tct ggt ggc gcc acc agt acc Asn Arg Gly Cys Lys Ala Met Phe Ile Ser Gly Ala Thr Ser Thr 275 280 285 | 864 |
| gtt ctt aag gac ggt atg acc cga gca cct gtt cggt ttc gct tcg Val Leu Lys Asp Gly Met Thr Arg Ala Pro Val Val Arg Phe Ala Ser 290 295 300 | 912 |
| gcg aga cga gct tcg gag ctt aag ttt ttc ttg gag aat cca gag aac Ala Arg Arg Ala Ser Glu Leu Lys Phe Phe Leu Glu Asn Pro Glu Asn 305 310 315 320 | 960 |
| ttt gat act ttg gca gta gtc ttc aac agg tcg agt aga ttt gca aga Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg 325 330 335 | 1008 |
| ctg caa agt gtt aaa tgc aca atc gcg ggg aag aat gct tat gta agg Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg 340 345 350 | 1056 |
| ttc tgt tgt agt act ggt gat gct atg ggg atg aat atg gtt tct aaa Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys 355 360 365 | 1104 |

| | |
|---|------|
| ggt gtg cag aat gtt ctt gag tat ctt acc gat gat ttc cct gac atg Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met 370 375 380 | 1152 |
| gat gtg att gga atc tct ggt aac ttc tgt tcg gac aag aaa cct gct Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala 385 390 395 400 | 1200 |
| gct gtg aac tgg att gag gga cgt ggt aaa tca gtt tgc gag gct Ala Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala 405 410 415 | 1248 |
| gta atc aga gga gag atc gtg aac aag gtc ttg aaa acg agc gtg gct Val Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala 420 425 430 | 1296 |
| gct tta gtc gag ctc aac atg ctc aag aac cta gct ggc tct gct gtt Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val 435 440 445 | 1344 |
| gca ggc tct cta ggt gga ttc aac gct cat gcc agt aac ata gtg tct Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser 450 455 460 | 1392 |
| gct gta ttc ata gct act ggc caa gat cca gct caa aac gtg gag agt Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser 465 470 475 480 | 1440 |
| tct caa tgc atc acc atg atg gaa gct att aat gac ggc aaa gat atc Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile 485 490 495 | 1488 |
| cat atc tca gtc act atg cca tct atc gag gtg ggg aca gtg gga gga His Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly 500 505 510 | 1536 |
| gga aca cag ctt gca tct caa tca gcg tgt tta aac ctg ctc gga gtt Gly Thr Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val 515 520 525 | 1584 |
| aaa gga gca agc aca gag tcg ccg gga atg aac gca agg agg cta gcg Lys Gly Ala Ser Thr Glu Ser Pro Gly Met Asn Ala Arg Arg Leu Ala 530 535 540 | 1632 |
| acg atc gta gcc gga gca gtt tta gct gga gag tta tct tta atg tca Thr Ile Val Ala Gly Ala Val Leu Ala Gly Glu Leu Ser Leu Met Ser 545 550 555 560 | 1680 |
| gca att gca gct gga cag ctt gtg aga agt cac atg aaa tac aat aga Ala Ile Ala Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg 565 570 575 | 1728 |
| tcc agc cga gac atc tct gga gca acg aca acg aca aca aca aca Ser Ser Arg Asp Ile Ser Gly Ala Thr Thr Thr Thr Thr Thr Thr 580 585 590 | 1776 |

tga 1779

<210> 54

<211> 592

<212> PRT

<213> Arabidopsis thaliana

<400> 54

Met Asp Leu Arg Arg Arg Pro Pro Lys Pro Pro Val Thr Asn Asn Asn
1 5 10 15

Asn Ser Asn Gly Ser Phe Arg Ser Tyr Gln Pro Arg Thr Ser Asp Asp
20 25 30

Asp His Arg Arg Arg Ala Thr Thr Ile Ala Pro Pro Pro Lys Ala Ser
35 40 45

Asp Ala Leu Pro Leu Pro Leu Tyr Leu Thr Asn Ala Val Phe Phe Thr
50 55 60

Leu Phe Phe Ser Val Ala Tyr Tyr Leu Leu His Arg Trp Arg Asp Lys
65 70 75 80

Ile Arg Tyr Asn Thr Pro Leu His Val Val Thr Ile Thr Glu Leu Gly
85 90 95

Ala Ile Ile Ala Leu Ile Ala Ser Phe Ile Tyr Leu Leu Gly Phe Phe
100 105 110

Gly Ile Asp Phe Val Gln Ser Phe Ile Ser Arg Ala Ser Gly Asp Ala
115 120 125

Trp Asp Leu Ala Asp Thr Ile Asp Asp Asp His Arg Leu Val Thr
130 135 140

Cys Ser Pro Pro Thr Pro Ile Val Ser Val Ala Lys Leu Pro Asn Pro
145 150 155 160

Glu Pro Ile Val Thr Glu Ser Leu Pro Glu Glu Asp Glu Glu Ile Val
165 170 175

Lys Ser Val Ile Asp Gly Val Ile Pro Ser Tyr Ser Leu Glu Ser Arg
180 185 190

Leu Gly Asp Cys Lys Arg Ala Ala Ser Ile Arg Arg Glu Ala Leu Gln
195 200 205

Arg Val Thr Gly Arg Ser Ile Glu Gly Leu Pro Leu Asp Gly Phe Asp
210 215 220

Tyr Glu Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Val Gly Tyr Ile
225 230 235 240

Gln Ile Pro Val Gly Ile Ala Gly Pro Leu Leu Leu Asp Gly Tyr Glu
245 250 255

Tyr Ser Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr
260 265 270

Asn Arg Gly Cys Lys Ala Met Phe Ile Ser Gly Gly Ala Thr Ser Thr
275 280 285

Val Leu Lys Asp Gly Met Thr Arg Ala Pro Val Val Arg Phe Ala Ser
290 295 300

Ala Arg Arg Ala Ser Glu Leu Lys Phe Phe Leu Glu Asn Pro Glu Asn
305 310 315 320

Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg
325 330 335

Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg
340 345 350

Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys
355 360 365

Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met
370 375 380

Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala
385 390 395 400

Ala Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala
405 410 415

Val Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala
420 425 430

Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val
435 440 445

Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser
450 455 460

Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser
465 470 475 480

Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile
485 490 495

His Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly
500 505 510

Gly Thr Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val
515 520 525

Lys Gly Ala Ser Thr Glu Ser Pro Gly Met Asn Ala Arg Arg Leu Ala
530 535 540

Thr Ile Val Ala Gly Ala Val Leu Ala Gly Glu Leu Ser Leu Met Ser
545 550 555 560

Ala Ile Ala Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg
565 570 575

Ser Ser Arg Asp Ile Ser Gly Ala Thr Thr Thr Thr Thr Thr Thr
580 585 590

<210> 55

<211> 1401

<212> DNA

<213> Arabidopsis thaliana ISPH

<220>

<221> CDS

<222> (1)..(1401)

<223>

| | | |
|---|--|-----|
| <400> 55 | | |
| atg gct gtt gcg ctc caa ttc agc cga tta tgc gtt cga ccg gat act | | 48 |
| Met Ala Val Ala Leu Gln Phe Ser Arg Leu Cys Val Arg Pro Asp Thr | | |
| 1 5 10 15 | | |
| ttc gtg cggtt gat cat ctc tct gga tcc gga tct ctc cgc cgc cgg | | 96 |
| Phe Val Arg Glu Asn His Leu Ser Gly Ser Gly Ser Leu Arg Arg Arg | | |
| 20 25 30 | | |
| aaa gct tta tca gtc cgg tgc tcg tct ggc gat gag aac gct cct tcg | | 144 |
| Lys Ala Leu Ser Val Arg Cys Ser Ser Gly Asp Glu Asn Ala Pro Ser | | |
| 35 40 45 | | |
| cca tcg gtg gtg atg gac tcc gat ttc gac gcc aag gtg ttc cgt aag | | 192 |
| Pro Ser Val Val Met Asp Ser Asp Phe Asp Ala Lys Val Phe Arg Lys | | |
| 50 55 60 | | |
| aac ttg acg aga agc gat aat tac aat cgt aaa ggg ttc ggt cat aag | | 240 |
| Asn Leu Thr Arg Ser Asp Asn Tyr Asn Arg Lys Gly Phe Gly His Lys | | |
| 65 70 75 80 | | |
| gag gag aca ctc aag ctc atg aat cga gag tac acc agt gat ata ttg | | 288 |
| Glu Glu Thr Leu Lys Leu Met Asn Arg Glu Tyr Thr Ser Asp Ile Leu | | |
| 85 90 95 | | |
| gag aca ctg aaa aca aat ggg tat act tat tct tgg gga gat gtt act | | 336 |
| Glu Thr Leu Lys Thr Asn Gly Tyr Thr Ser Trp Gly Asp Val Thr | | |
| 100 105 110 | | |
| gtg aaa ctc gct aaa gca tat ggt ttt tgc tgg ggt gtt gag cgt gct | | 384 |
| Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala | | |
| 115 120 125 | | |
| gtt cag att gca tat gaa gca cga aag cag ttt cca gag gag agg ctt | | 432 |
| Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu | | |
| 130 135 140 | | |
| tgg att act aac gaa atc att cat aac ccg acc gtc aat aag agg ttg | | 480 |
| Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu | | |
| 145 150 155 160 | | |

| | |
|---|------|
| gaa gat atg gat gtt aaa att att ccg gtt gag gat tca aag aaa cag Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln 165 170 175 | 528 |
| ttt gat gta gta gag aaa gat gat gtg gtt atc ctt cct gcg ttt gga Phe Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly 180 185 190 | 576 |
| gct ggt gtt gac gag atg tat gtt ctt aat gat aaa aag gtg caa att Ala Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile 195 200 205 | 624 |
| gtt gac acg act tgt cct tgg gtg aca aag gtc tgg aac acg gtt gag Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu 210 215 220 | 672 |
| aag cac aag aag ggg gaa tac aca tca gta atc cat ggt aaa tat aat Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn 225 230 235 240 | 720 |
| cat gaa gag acg att gca act gcg tct ttt gca gga aag tac atc att His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile 245 250 255 | 768 |
| gta aag aac atg aaa gag gca aat tac gtt tgt gat tac att ctc ggt Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly 260 265 270 | 816 |
| ggc caa tac gat gga tct agc tcc aca aaa gag gag ttc atg gag aaa Gly Gln Tyr Asp Gly Ser Ser Thr Lys Glu Glu Phe Met Glu Lys 275 280 285 | 864 |
| ttc aaa tac gca att tcg aag ggt ttc gat ccc gac aat gac ctt gtc Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val 290 295 300 | 912 |
| aaa gtt ggt att gca aac caa aca acg atg cta aag gga gaa aca gag Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu 305 310 315 320 | 960 |
| gag ata gga aga tta ctc gag aca aca atg atg cgc aag tat gga gtg Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val 325 330 335 | 1008 |
| gaa aat gta agc gga cat ttc atc agc ttc aac aca ata tgc gac gct Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala 340 345 350 | 1056 |
| actcaa gag cga caa gac gca atc tat gag cta gtg gaa gag aag att Thr Gln Glu Arg Gln Asp Ala Ile Tyr Glu Leu Val Glu Glu Lys Ile 355 360 365 | 1104 |
| gac ctc atg cta gtg gtt ggc gga tgg aat tca agt aac acc tct cac Asp Leu Met Leu Val Val Gly Gly Trp Asn Ser Ser Asn Thr Ser His 370 375 380 | 1152 |

ctt cag gaa atc tca gag gca cg^g gga atc cca tct tac tgg atc gat 1200
 Leu Gln Glu Ile Ser Glu Ala Arg Gly Ile Pro Ser Tyr Trp Ile Asp
 385 390 395 400

agt gag aaa cgg ata gga cct ggg aat aaa ata gcc tat aag ctc cac 1248
 Ser Glu Lys Arg Ile Gly Pro Gly Asn Lys Ile Ala Tyr Lys Leu His
 405 410 415

tat gga gaa ctg gtc gag aag gaa aac ttt ctc cca aag gga cca ata 1296
Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile
420 425 430

aca atc ggt gtg aca tca ggt gca tca acc ccg gat aag gtc gtg gaa 1344
 Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu
 435 440 445

gat gct ttg gtg aag gtg ttc gac att aaa cgt gaa gag tta ttg cag 1392
Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln
450 455 . 460

ctg gct tga 1401
Leu Ala
465

<210> 56

<211> 466

<212> PRT

<213> *Arabidopsis thaliana* ISPH

<400> 56

Met Ala Val Ala Leu Gln Phe Ser Arg Leu Cys Val Arg Pro Asp Thr
1 5 10 15

Phe Val Arg Glu Asn His Leu Ser Gly Ser Gly Ser Leu Arg Arg Arg
 20 25 . 30

Lys Ala Leu Ser Val Arg Cys Ser Ser Gly Asp Glu Asn Ala Pro Ser
35 40 45

Pro Ser Val Val Met Asp Ser Asp Phe Asp Ala Lys Val Phe Arg Lys
50 55 60

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Leu | Thr | Arg | Ser | Asp | Asn | Tyr | Asn | Arg | Lys | Gly | Phe | Gly | His | Lys |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |

Glu Glu Thr Leu Lys Leu Met Asn Arg Glu Tyr Thr Ser Asp Ile Leu
85 90 95

Glu Thr Leu Lys Thr Asn Gly Tyr Thr Tyr Ser Trp Gly Asp Val Thr
100 105 110

Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala
115 120 125

Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu
130 135 140

Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu
145 150 155 160

Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln
165 170 175

Phe Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly
180 185 190

Ala Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile
195 200 205

Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu
210 215 220

Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn
225 230 235 240

His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile
245 250 255

Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly
260 265 270

Gly Gln Tyr Asp Gly Ser Ser Ser Thr Lys Glu Glu Phe Met Glu Lys
275 280 285

Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val
290 295 300

Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu.
305 310 315 320

Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val
325 330 335

Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala
340 345 350

Thr Gln Glu Arg Gln Asp Ala Ile Tyr Glu Leu Val Glu Glu Lys Ile
355 360 365

Asp Leu Met Leu Val Val Gly Gly Trp Asn Ser Ser Asn Thr Ser His
370 375 380

Leu Gln Glu Ile Ser Glu Ala Arg Gly Ile Pro Ser Tyr Trp Ile Asp
385 390 395 400

Ser Glu Lys Arg Ile Gly Pro Gly Asn Lys Ile Ala Tyr Lys Leu His
405 410 415

Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile
420 425 430

Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu
435 440 445

Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln
450 455 460

Leu Ala
465

<210> 57

<211> 2160

<212> DNA

<213> Lycopersicon esculentum

100

<220>

<221> CDS

<222> (1) .. (2160)

<223>

| | | | |
|---|-----|--|--|
| <400> 57 | | | |
| atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt | 48 | | |
| Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly | | | |
| 1 5 10 15 | | | |
| gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att | 96 | | |
| Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile | | | |
| 20 25 30 | | | |
| cat gga aca gat ctg cag ttt ttg ttc caa cac aag ctt act cat gag | 144 | | |
| His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu | | | |
| 35 40 45 | | | |
| gtc aag aaa agg tca cgt gtg gtt cag gct tcc tta tca gaa tct gga | 192 | | |
| Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly | | | |
| 50 55 60 | | | |
| gaa tac tac aca cag aga ccg cca acg cct att ttg gac act gtg aac | 240 | | |
| Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn | | | |
| 65 70 75 80 | | | |
| tat ccc att cat atg aaa aat ctg tct ctg aag gaa ctt aaa caa cta | 288 | | |
| Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu | | | |
| 85 90 95 | | | |
| gca gat gaa cta agg tca gat aca att ttc aat gta tca aag act ggg | 336 | | |
| Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly | | | |
| 100 105 110 | | | |
| ggt cac ctt ggc tca agt ctt ggt gtt gag ctg act gtt gct ctt | 384 | | |
| Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu | | | |
| 115 120 125 | | | |
| cat tat gtc ttc aat gca ccg caa gat agg att ctc tgg gat gtt ggt | 432 | | |
| His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly | | | |
| 130 135 140 | | | |
| cat cag tct tat cct cac aaa atc ttg act ggt aga agg gac aag atg | 480 | | |
| His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met | | | |
| 145 150 155 160 | | | |
| tcg aca tta agg cag aca gat ggt ctt gca gga ttt act aag cga tcg | 528 | | |
| Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser | | | |
| 165 170 175 | | | |
| gag agt gaa tat gat tgc ttt ggc acc ggc cac agt tcc acc acc atc | 576 | | |

101

| | | |
|---|-----|------|
| Glu Ser Glu Tyr Asp Cys Phe Gly Thr Gly His Ser Ser Thr Thr Ile | | |
| 180 | 185 | 190 |
| tca gca ggc cta ggg atg gct gtt ggt aga gat cta aaa gga aga aac | | 624 |
| Ser Ala Gly Leu Gly Met Ala Val Gly Arg Asp Leu Lys Gly Arg Asn | | |
| 195 | 200 | 205 |
| aac aat gtt att gcc gta ata ggt gat ggt gcc atg aca gca ggt caa | | 672 |
| Asn Asn Val Ile Ala Val Ile Gly Asp Gly Ala Met Thr Ala Gly Gln | | |
| 210 | 215 | 220 |
| gct tat gaa gcc atg aat aat gct ggt tac ctg gac tct gac atg att | | 720 |
| Ala Tyr Glu Ala Met Asn Asn Ala Gly Tyr Leu Asp Ser Asp Met Ile | | |
| 225 | 230 | 235 |
| gtt atc tta aac gac aat aga caa gtt tct tta cct act gct act ctg | | 768 |
| Val Ile Leu Asn Asp Asn Arg Gln Val Ser Leu Pro Thr Ala Thr Leu | | |
| 245 | 250 | 255 |
| gat ggg cca gtt gct cct gtt gga gct cta agt agt gct ttg agc agg | | 816 |
| Asp Gly Pro Val Ala Pro Val Gly Ala Leu Ser Ser Ala Leu Ser Arg | | |
| 260 | 265 | 270 |
| tta cag tct aat agg cct ctc aga gaa cta aga gaa gtc gca aag gga | | 864 |
| Leu Gln Ser Asn Arg Pro Leu Arg Glu Leu Arg Glu Val Ala Lys Gly | | |
| 275 | 280 | 285 |
| gtt act aag cag att ggt ggt cct atg cat gag ctt gct gca aaa gtt | | 912 |
| Val Thr Lys Gln Ile Gly Gly Pro Met His Glu Leu Ala Ala Lys Val | | |
| 290 | 295 | 300 |
| gat gaa tat gct cgt ggc atg att agt ggt tct gga tca aca ttg ttt | | 960 |
| Asp Glu Tyr Ala Arg Gly Met Ile Ser Gly Ser Gly Ser Thr Leu Phe | | |
| 305 | 310 | 315 |
| 320 | | |
| gaa gaa ctt gga ctt tac tat att ggt cct gtg gat ggt cac aac att | | 1008 |
| Glu Glu Leu Gly Leu Tyr Ile Gly Pro Val Asp Gly His Asn Ile | | |
| 325 | 330 | 335 |
| gat gat cta att gcg att ctc aaa gag gtt aga agt act aaa aca aca | | 1056 |
| Asp Asp Leu Ile Ala Ile Leu Lys Glu Val Arg Ser Thr Lys Thr Thr | | |
| 340 | 345 | 350 |
| ggt cca gta ctg atc cat gtt gtc act gag aaa ggc aga ggt tat cca | | 1104 |
| Gly Pro Val Leu Ile His Val Val Thr Glu Lys Gly Arg Gly Tyr Pro | | |
| 355 | 360 | 365 |
| tat gct gag aga gct gca gat aag tat cat gga gtt gcc aag ttt gat | | 1152 |
| Tyr Ala Glu Arg Ala Ala Asp Lys Tyr His Gly Val Ala Lys Phe Asp | | |
| 370 | 375 | 380 |
| cca gca aca gga aag caa ttc aaa gcc agt gcc aag aca cag tcc tat | | 1200 |
| Pro Ala Thr Gly Lys Gln Phe Lys Ala Ser Ala Lys Thr Gln Ser Tyr | | |
| 385 | 390 | 395 |
| 400 | | |
| aca aca tat ttt gcc gag gct tta att gca gaa gca gaa gca gat aaa | | 1248 |

102

| | | | | |
|---|-----|-----|-----|------|
| Thr Thr Tyr Phe Ala Glu Ala Leu Ile Ala Glu Ala Asp Lys | | | | |
| 405 | 410 | 415 | | |
| gac att gtt gca atc cat gct gcc atg ggg ggt ggg acc gga atg aac | | | | 1296 |
| Asp Ile Val Ala Ile His Ala Ala Met Gly Gly Gly Thr Gly Met Asn | | | | |
| 420 | 425 | 430 | | |
| ctt ttc cat cgt cgc ttc cca aca agg tgt ttt gat gtt gga ata gca | | | | 1344 |
| Leu Phe His Arg Arg Phe Pro Thr Arg Cys Phe Asp Val Gly Ile Ala | | | | |
| 435 | 440 | 445 | | |
| gaa caa cat gca gta acc ttt gct gct gga ttg gct tgt gaa ggc att | | | | 1392 |
| Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Cys Glu Gly Ile | | | | |
| 450 | 455 | 460 | | |
| aaa cct ttc tgt gca atc tat tcg tct ttc atg cag agg gct tat gac | | | | 1440 |
| Lys Pro Phe Cys Ala Ile Tyr Ser Ser Phe Met Gln Arg Ala Tyr Asp | | | | |
| 465 | 470 | 475 | 480 | |
| cag gta gtg cat gac gtt gat ttg caa aag ctg ccc gtg agg ttt gca | | | | 1488 |
| Gln Val Val His Asp Val Asp Leu Gln Lys Leu Pro Val Arg Phe Ala | | | | |
| 485 | 490 | 495 | | |
| atg gac aga gca ggt ctt gtt gga gca gat ggt cca aca cat tgt ggt | | | | 1536 |
| Met Asp Arg Ala Gly Leu Val Gly Ala Asp Gly Pro Thr His Cys Gly | | | | |
| 500 | 505 | 510 | | |
| gca ttt gat gtt act tac atg gca tgt ctt cct aac atg gtt gta atg | | | | 1584 |
| Ala Phe Asp Val Thr Tyr Met Ala Cys Leu Pro Asn Met Val Val Met | | | | |
| 515 | 520 | 525 | | |
| gct cct tct gat gaa gcg gag cta ttt cac atg gta gca act gct gcc | | | | 1632 |
| Ala Pro Ser Asp Glu Ala Glu Leu Phe His Met Val Ala Thr Ala Ala | | | | |
| 530 | 535 | 540 | | |
| gcc att gat gac aga cca agt tgt ttt aga tac cca aga gga aat ggg | | | | 1680 |
| Ala Ile Asp Asp Arg Pro Ser Cys Phe Arg Tyr Pro Arg Gly Asn Gly | | | | |
| 545 | 550 | 555 | 560 | |
| atc ggt gta gag ctt ccg gct gga aac aaa gga att cct ctt gag gtt | | | | 1728 |
| Ile Gly Val Glu Leu Pro Ala Gly Asn Lys Gly Ile Pro Leu Glu Val | | | | |
| 565 | 570 | 575 | | |
| ggtaaa ggt agg ata ttg att gag ggg gag aga gtg gct cta ttg gga | | | | 1776 |
| Gly Lys Gly Arg Ile Leu Ile Glu Gly Glu Arg Val Ala Leu Leu Gly | | | | |
| 580 | 585 | 590 | | |
| tat ggc tca gca gtg cag aac tgt ttg gat gca gat gca cgt ttc tgc aaa cca | | | | 1824 |
| Tyr Gly Ser Ala Val Gln Asn Cys Leu Asp Ala Ala Ile Val Leu Glu | | | | |
| 595 | 600 | 605 | | |
| tcc cgc ggc tta caa gta aca gtt gca gat gca cgt ttc tgc aaa cca | | | | 1872 |
| Ser Arg Gly Leu Gln Val Thr Val Ala Asp Ala Arg Phe Cys Lys Pro | | | | |
| 610 | 615 | 620 | | |
| ctg gac cat gcc ctc ata agg agc ctt gca aaa tca cat gaa gtg cta | | | | 1920 |

103

| | | | |
|---|-----|-----|------|
| Leu Asp His Ala Leu Ile Arg Ser Leu Ala Lys Ser His Glu Val Leu | | | |
| 625 | 630 | 635 | 640 |
| atc act gtc gaa gaa gga tca att gga ggt ttt gga tct cat gtt gtt | | | 1968 |
| Ile Thr Val Glu Glu Gly Ser Ile Gly Gly Phe Gly Ser His Val Val | | | |
| 645 | 650 | 655 | |
| cag ttc atg gcc tta gat ggg ctt ctt gat ggc aag ttg aag tgg aga | | | 2016 |
| Gln Phe Met Ala Leu Asp Gly Leu Leu Asp Gly Lys Leu Lys Trp Arg | | | |
| 660 | 665 | 670 | |
| cca ata gtt ctt cct gat cga tac att gac cat gga tct cct gtt gat | | | 2064 |
| Pro Ile Val Leu Pro Asp Arg Tyr Ile Asp His Gly Ser Pro Val Asp | | | |
| 675 | 680 | 685 | |
| cag ttg gcg gaa gct ggc cta aca cca tct cac att gca gca aca gta | | | 2112 |
| Gln Leu Ala Glu Ala Gly Leu Thr Pro Ser His Ile Ala Ala Thr Val | | | |
| 690 | 695 | 700 | |
| ttt aac ata ctt gga caa acc aga gag gct cta gag gtc atg aca taa | | | 2160 |
| Phe Asn Ile Leu Gly Gln Thr Arg Glu Ala Leu Glu Val Met Thr | | | |
| 705 | 710 | 715 | |
| <210> 58 | | | |
| <211> 719 | | | |
| <212> PRT | | | |
| <213> Lycopersicon esculentum | | | |
| <400> 58 | | | |
| Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly | | | |
| 1 | 5 | 10 | 15 |
| Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile | | | |
| 20 | 25 | 30 | |
| His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu | | | |
| 35 | 40 | 45 | |
| Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly | | | |
| 50 | 55 | 60 | |
| Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn | | | |
| 65 | 70 | 75 | 80 |

104

Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu
85 90 95

Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly
100 105 110

Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu
115 120 125

His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly
130 135 140

His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met
145 150 155 160

Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser
165 170 175

Glu Ser Glu Tyr Asp Cys Phe Gly Thr Gly His Ser Ser Thr Thr Ile
180 185 190

Ser Ala Gly Leu Gly Met Ala Val Gly Arg Asp Leu Lys Gly Arg Asn
195 200 205

Asn Asn Val Ile Ala Val Ile Gly Asp Gly Ala Met Thr Ala Gly Gln
210 215 220

Ala Tyr Glu Ala Met Asn Asn Ala Gly Tyr Leu Asp Ser Asp Met Ile
225 230 235 240

Val Ile Leu Asn Asp Asn Arg Gln Val Ser Leu Pro Thr Ala Thr Leu
245 250 255

Asp Gly Pro Val Ala Pro Val Gly Ala Leu Ser Ser Ala Leu Ser Arg
260 265 270

Leu Gln Ser Asn Arg Pro Leu Arg Glu Leu Arg Glu Val Ala Lys Gly
275 280 285

Val Thr Lys Gln Ile Gly Gly Pro Met His Glu Leu Ala Ala Lys Val
290 295 300

105

Asp Glu Tyr Ala Arg Gly Met Ile Ser Gly Ser Gly Ser Thr Leu Phe
305 310 315 320

Glu Glu Leu Gly Leu Tyr Tyr Ile Gly Pro Val Asp Gly His Asn Ile
325 330 335

Asp Asp Leu Ile Ala Ile Leu Lys Glu Val Arg Ser Thr Lys Thr Thr
340 345 350

Gly Pro Val Leu Ile His Val Val Thr Glu Lys Gly Arg Gly Tyr Pro
355 360 365

Tyr Ala Glu Arg Ala Ala Asp Lys Tyr His Gly Val Ala Lys Phe Asp
370 375 380

Pro Ala Thr Gly Lys Gln Phe Lys Ala Ser Ala Lys Thr Gln Ser Tyr
385 390 395 400

Thr Thr Tyr Phe Ala Glu Ala Leu Ile Ala Glu Ala Glu Ala Asp Lys
405 410 415

Asp Ile Val Ala Ile His Ala Ala Met Gly Gly Gly Thr Gly Met Asn
420 425 430

Leu Phe His Arg Arg Phe Pro Thr Arg Cys Phe Asp Val Gly Ile Ala
435 440 445

Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Cys Glu Gly Ile
450 455 460

Lys Pro Phe Cys Ala Ile Tyr Ser Ser Phe Met Gln Arg Ala Tyr Asp
465 470 475 480

Gln Val Val His Asp Val Asp Leu Gln Lys Leu Pro Val Arg Phe Ala
485 490 495

Met Asp Arg Ala Gly Leu Val Gly Ala Asp Gly Pro Thr His Cys Gly
500 505 510

Ala Phe Asp Val Thr Tyr Met Ala Cys Leu Pro Asn Met Val Val Met
515 520 525

106

Ala Pro Ser Asp Glu Ala Glu Leu Phe His Met Val Ala Thr Ala Ala
530 535 540

Ala Ile Asp Asp Arg Pro Ser Cys Phe Arg Tyr Pro Arg Gly Asn Gly
545 550 555 560

Ile Gly Val Glu Leu Pro Ala Gly Asn Lys Gly Ile Pro Leu Glu Val
565 570 575

Gly Lys Gly Arg Ile Leu Ile Glu Gly Glu Arg Val Ala Leu Leu Gly
580 585 590

Tyr Gly Ser Ala Val Gln Asn Cys Leu Asp Ala Ala Ile Val Leu Glu
595 600 605

Ser Arg Gly Leu Gln Val Thr Val Ala Asp Ala Arg Phe Cys Lys Pro
610 615 620

Leu Asp His Ala Leu Ile Arg Ser Leu Ala Lys Ser His Glu Val Leu
625 630 635 640

Ile Thr Val Glu Glu Gly Ser Ile Gly Gly Phe Gly Ser His Val Val
645 650 655

Gln Phe Met Ala Leu Asp Gly Leu Leu Asp Gly Lys Leu Lys Trp Arg
660 665 670

Pro Ile Val Leu Pro Asp Arg Tyr Ile Asp His Gly Ser Pro Val Asp
675 680 685

Gln Leu Ala Glu Ala Gly Leu Thr Pro Ser His Ile Ala Ala Thr Val
690 695 700

Phe Asn Ile Leu Gly Gln Thr Arg Glu Ala Leu Glu Val Met Thr
705 710 715

<210> 59

<211> 1434

<212> DNA

<213> *Arabidopsis thaliana*

<220>

<221> CDS

<222> (1) .. (1434) |

<223> | - |

<400> 59

atg atg aca tta aac tca cta tct cca gct gaa tcc aaa gct att tct 48
 Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser
 1 5 10 15

| |
 ttc ttg gat acc tcc agg ttc aat cca atc cct aaa ctc tca ggt ggg 96
 Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly
 20 25 30

ttt agt ttg agg agg aat caa ggg aga ggt ttt gga aaa ggt gtt 144
 Phe Ser Leu Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val
 35 40 45

| |
 aag tgt tca gtg aaa gtg cag cag caa caa caa cct cct cca gca tgg 192
 Lys Cys Ser Val Lys Val Gln Gln Gln Gln Pro Pro Pro Ala Trp
 50 55 60

cct ggg aga gct gtc cct gag gcg cct cgt caa tct tgg gat gga cca 240
 Pro Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro
 65 70 75 80

aaa ccc atc tct atc gtt gga tct act ggt tct att ggc act cag aca 288
 Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr
 85 90 95

ttg gat att gtg gct gag aat cct gac aaa ttc aga gtt gtg gct cta 336
 Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu
 100 105 110

gct gct ggt tcg aat gtt act cta ctt gct gat cag gta agg aga ttt 384
 Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe
 115 120 125

aag cct gca ttg gtt gct gtt aga aac gag tca ctg att aat gag ctt 432
 Lys Pro Ala Leu Val Ala Val Arg Asn Glu Ser Leu Ile Asn Glu Leu
 130 135 140

aaa gag gct tta gct gat ttg gac tat aaa ctc gag att att cca gga 480
 Lys Glu Ala Leu Ala Asp Leu Asp Tyr Lys Leu Glu Ile Ile Pro Gly
 145 150 155 160

| |
 gag caa gga gtg att gag gtt gcc cga cat cct gaa gct gta acc gtt 528
 Glu Gln Gly Val Ile Glu Val Ala Arg His Pro Glu Ala Val Thr Val
 165 170 175

| | | | |
|---|-----|------|-----|
| gtt acc gga ata gta ggt tgt gcg gga cta aag cct acg gtt gct gca | | 576 | |
| Val Thr Gly Ile Val Gly Cys Ala Gly Leu Lys Pro Thr Val Ala Ala | | | |
| 180 | 185 | 190 | |
| att gaa gca gga aag gac att gct ctt gca aac aaa gag aca tta atc | | 624 | |
| Ile Glu Ala Gly Lys Asp Ile Ala Leu Ala Asn Lys Glu Thr Leu Ile | | | |
| 195 | 200 | 205 | |
| gca ggt ggt cct ttc gtg ctt ccg ctt gcc aac aaa cat aat gta aag | | 672 | |
| Ala Gly Gly Pro Phe Val Leu Pro Leu Ala Asn Lys His Asn Val Lys | | | |
| 210 | 215 | 220 | |
| att ctt ccg gca gat tca gaa cat tct gcc ata ttt cag tgt att caa | | 720 | |
| Ile Leu Pro Ala Asp Ser Glu His Ser Ala Ile Phe Gln Cys Ile Gln | | | |
| 225 | 230 | 235 | 240 |
| ggg ttg cct gaa ggc gct ctg cgc aag ata atc ttg act gca tct ggt | | 768 | |
| Gly Leu Pro Glu Gly Ala Leu Arg Lys Ile Ile Leu Thr Ala Ser Gly | | | |
| 245 | 250 | 255 | |
| gga gct ttt agg gat tgg cct gtc gaa aag cta aag gaa gtt aaa gta | | 816 | |
| Gly Ala Phe Arg Asp Trp Pro Val Glu Lys Leu Lys Glu Val Lys Val | | | |
| 260 | 265 | 270 | |
| gcg gat gcg ttg aag cat cca aac tgg aac atg gga aag aaa atc act | | 864 | |
| Ala Asp Ala Leu Lys His Pro Asn Trp Asn Met Gly Lys Lys Ile Thr | | | |
| 275 | 280 | 285 | |
| gtg gac tct gct acg ctt ttc aac aag ggt ctt gag gtc att gaa gcg | | 912 | |
| Val Asp Ser Ala Thr Leu Phe Asn Lys Gly Leu Glu Val Ile Glu Ala | | | |
| 290 | 295 | 300 | |
| cat tat ttg ttt gga gct gag tat gac gat ata gag att gtc att cat | | 960 | |
| His Tyr Leu Phe Gly Ala Glu Tyr Asp Asp Ile Glu Ile Val Ile His | | | |
| 305 | 310 | 315 | 320 |
| ccg caa agt atc ata cat tcc atg att gaa aca cag gat tca tct gtg | | 1008 | |
| Pro Gln Ser Ile Ile His Ser Met Ile Glu Thr Gln Asp Ser Ser Val | | | |
| 325 | 330 | 335 | |
| ctt gct caa ttg ggt tgg cct gat atg cgt tta ccg att ctc tac acc | | 1056 | |
| Leu Ala Gln Leu Gly Trp Pro Asp Met Arg Leu Pro Ile Leu Tyr Thr | | | |
| 340 | 345 | 350 | |
| atg tca tgg ccc gat aga gtt cct tgt tct gaa gta act tgg cca aga | | 1104 | |
| Met Ser Trp Pro Asp Arg Val Pro Cys Ser Glu Val Thr Trp Pro Arg | | | |
| 355 | 360 | 365 | |
| ctt gac ctt tgc aaa ctc ggt tca ttg act ttc aag aaa cca gac aat | | 1152 | |
| Leu Asp Leu Cys Lys Leu Gly Ser Leu Thr Phe Lys Lys Pro Asp Asn | | | |
| 370 | 375 | 380 | |
| gtg aaa tac cca tcc atg gat ctt gct tat gct gct gga cga gct gga | | 1200 | |
| Val Lys Tyr Pro Ser Met Asp Leu Ala Tyr Ala Ala Gly Arg Ala Gly | | | |
| 385 | 390 | 395 | 400 |

```

ggc aca atg act gga gtt ctc agc gcc gcc aat gag aaa gct gtt gaa      1248
Gly Thr Met Thr Gly Val Leu Ser Ala Ala Asn Glu Lys Ala Val Glu
        405           410           415

```

```

atg ttc att gat gaa aag ata agc tat ttg gat atc ttc aag gtt gtg      1296
Met Phe Ile Asp Glu Lys Ile Ser Tyr Leu Asp Ile Phe Lys Val Val
        420           425           430

```

gaa tta aca tgc gat aaa cat cga aac gag ttg gta aca tca ccg tct 1344
 Glu Leu Thr Cys Asp Lys His Arg Asn Glu Leu Val Thr Ser Pro Ser
 435 440 445

ctt gaa gag att gtt cac tat gac ttg tgg gca cgt gaa tat gcc gcg 1392
Leu Glu Glu Ile Val His Tyr Asp Leu Trp Ala Arg Glu Tyr Ala Ala
450 . 455 . 460

```

aat gtg cag ctt tct tct ggt gct agg cca gtt cat gca tga      1434
Asn Val Gln Leu Ser Ser Gly Ala Arg Pro Val His Ala
465           470           475

```

<210> 60

<211> 477

<212> PRT

<213> *Arabidopsis thaliana*

<400> 60

```

Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser
1           5           10          15

```

Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly
20 25 30

Phe Ser Leu Arg Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val
 35 40 45

Lys ·Cys Ser Val Lys Val Gln Gln Gln Gln Gln Pro Pro Pro Ala Trp
 50 55 60

Pro Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro
65 70 75 80

Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr
85 90 95

Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu
100 105 110

Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe
115 120 125

Lys Pro Ala Leu Val Ala Val Arg Asn Glu Ser Leu Ile Asn Glu Leu
130 135 140

Lys Glu Ala Leu Ala Asp Leu Asp Tyr Lys Leu Glu Ile Ile Pro Gly
145 150 155 160

Glu Gln Gly Val Ile Glu Val Ala Arg His Pro Glu Ala Val Thr Val
165 170 175

Val Thr Gly Ile Val Gly Cys Ala Gly Leu Lys Pro Thr Val Ala Ala
180 185 190

Ile Glu Ala Gly Lys Asp Ile Ala Leu Ala Asn Lys Glu Thr Leu Ile
195 200 205

Ala Gly Gly Pro Phe Val Leu Pro Leu Ala Asn Lys His Asn Val Lys
210 215 220

Ile Leu Pro Ala Asp Ser Glu His Ser Ala Ile Phe Gln Cys Ile Gln
225 230 235 240

Gly Leu Pro Glu Gly Ala Leu Arg Lys Ile Ile Leu Thr Ala Ser Gly
245 250 255

Gly Ala Phe Arg Asp Trp Pro Val Glu Lys Leu Lys Glu Val Lys Val
260 265 270

Ala Asp Ala Leu Lys His Pro Asn Trp Asn Met Gly Lys Lys Ile Thr
275 280 285

Val Asp Ser Ala Thr Leu Phe Asn Lys Gly Leu Glu Val Ile Glu Ala
290 295 300

His Tyr Leu Phe Gly Ala Glu Tyr Asp Asp Ile Glu Ile Val Ile His
305 310 315 320

Pro Gln Ser Ile Ile His Ser Met Ile Glu Thr Gln Asp Ser Ser Val
325 330 335

Leu Ala Gln Leu Gly Trp Pro Asp Met Arg Leu Pro Ile Leu Tyr Thr
340 345 350

Met Ser Trp Pro Asp Arg Val Pro Cys Ser Glu Val Thr Trp Pro Arg
355 360 365

Leu Asp Leu Cys Lys Leu Gly Ser Leu Thr Phe Lys Lys Pro Asp Asn
370 375 380

Val Lys Tyr Pro Ser Met Asp Leu Ala Tyr Ala Ala Gly Arg Ala Gly
385 390 395 400

Gly Thr Met Thr Gly Val Leu Ser Ala Ala Asn Glu Lys Ala Val Glu
405 410 415

Met Phe Ile Asp Glu Lys Ile Ser Tyr Leu Asp Ile Phe Lys Val Val
420 425 430

Glu Leu Thr Cys Asp Lys His Arg Asn Glu Leu Val Thr Ser Pro Ser
435 440 445

Leu Glu Glu Ile Val His Tyr Asp Leu Trp Ala Arg Glu Tyr Ala Ala
450 455 460

Asn Val Gln Leu Ser Ser Gly Ala Arg Pro Val His Ala
465 470 475

<210> 61

<211> 884

<212> DNA

<213> Adonis palaestina clone ApIPI28

<220>

<221> CDS

<222> (180)...(884)

<223>

| | | | | |
|---|--|--|--|-----|
| <400> 61 | | | | |
| cgtcgatcag gattaatcct ttatatagtt tcttctccac caccactaaa acattatcag | | | | 60 |
| cttcgtgttc ttctccgcgtt gttcatcttc agcagcggtt tcgtacttctt tctatccctt | | | | 120 |
| cttccatcac taacagtccct cgccgagggt tgaatcggtt gttcgctca acgtcgact | | | | 179 |
| atg ggt gaa gtc gct gat gct ggt atg gat gcc gtc cag aag cgg ctt | | | | 227 |
| Met Gly Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu | | | | |
| 1 5 10 15 | | | | |
| atg ttc gac gat gaa tgt att ttg gtg gat gag aat gac aag gtc gtc | | | | 275 |
| Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val | | | | |
| 20 25 30 | | | | |
| gga cat gat tcc aaa tac aac tgt cat ttg atg gaa aag ata gag gca | | | | 323 |
| Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala | | | | |
| 35 40 45 | | | | |
| gaa aac ttg ctt cac aga gcc ttc agt gtt ttc tta ttc aac tca aaa | | | | 371 |
| Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys | | | | |
| 50 55 60 | | | | |
| tac gag ttg ctt ctt cag caa cga tct gca acg aag gta aca ttc ccg | | | | 419 |
| Tyr Glu Leu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro | | | | |
| 65 70 75 80 | | | | |
| ctc gta tgg aca aac acc tgt tgc agc cat ccc ctc ttc cgt gat tcc | | | | 467 |
| Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser | | | | |
| 85 90 95 | | | | |
| gaa ctc ata gaa gaa aat ttt ctc ggg gta cga aac gct gca caa agg | | | | 515 |
| Glu Leu Ile Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg | | | | |
| 100 105 110 | | | | |
| aag ctt tta gac gag cta ggc att cca gct gaa gac gta cca gtt gat | | | | 563 |
| Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp | | | | |
| 115 120 125 | | | | |
| gaa ttc act cct ctt ggt cgc att ctt tac aaa gct cca tct gac gga | | | | 611 |
| Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly | | | | |
| 130 135 140 | | | | |
| aaa tgg gga gag cac gaa ctg gac tat ctt ctg ttt att gtc cga gat | | | | 659 |
| Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp | | | | |
| 145 150 155 160 | | | | |
| gtg aaa tac gat cca aac cca gat gaa gtt gct gac gct aag tac gtt | | | | 707 |
| Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val | | | | |
| 165 170 175 | | | | |

aat cgc gag gag ttg aaa gag ata ctg aga aaa gct gat gca ggt gaa 755
Asn Arg Glu Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu
180 185 190

gag gga ata aag ttg tct cct tgg ttt aga ttg gtt gtg gat aac ttt 803
Glu Gly Ile Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe
195 200 205

ttg ttc aag tgg tgg gat cat gta gag gag ggg aag att aag gac gtc 851
Leu Phe Lys Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val
210 215 220

gcc gac atg aaa act atc cac aag ttg act taa 884
Ala Asp Met Lys Thr Ile His Lys Leu Thr
225 230

<210> 62

<211> 234

<212> PRT

<213> Adonis palaestina clone ApIPI28

<400> 62

Met Gly Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu
1 5 10 15

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val
20 25 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala
35 40 45

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro
65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser
85 90 95

Glu Leu Ile Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg
100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp
115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly
130 135 140

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp
145 150 155 160

Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val
165 170 175

Asn Arg Glu Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu
180 185 190

Glu Gly Ile Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe
195 [] 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val
210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr
225 230

<210> 63

<211> 1402

<212> DNA

<213> Arabidopsis thaliana []

<220>

<221> CDS

<222> (52)..(1317)

<223>

<400> 63
aagtctttgc ctctttgggt tactttcctc tgttttcgat ccatttagaa a atg tta 57

Met Leu
1

| | |
|--|-----|
| ttc acg agg agt gtt gct cgg att tct tct aag ttt ctg aga aac cgt Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg Asn Arg | 105 |
| 5 10 15 | |
| agc ttc tat ggc tcc tct caa tct ctc gcc tct cat cgg ttc gca atc Ser Phe Tyr Gly Ser Ser Gln Ser Leu Ala Ser His Arg Phe Ala Ile | 153 |
| 20 25 30 | |
| att ccc gat cag ggt cac tct tgt tct gac tct cca cac aag ggt tac Ile Pro Asp Gln Gly His Ser Cys Ser Asp Ser Pro His Lys Gly Tyr | 201 |
| 35 40 45 50 | |
| gtt tgc aga aca act tat tca ttg aaa tct ccg gtt ttt ggt gga ttt Val Cys Arg Thr Tyr Ser Leu Lys Ser Pro Val Phe Gly Gly Phe | 249 |
| 55 60 65 | |
| agt cat caa ctc tat cac cag agt agc tcc ttg gtt gag gag gag ctt Ser His Gln Leu Tyr His Gln Ser Ser Leu Val Glu Glu Glu Leu | 297 |
| 70 75 80 | |
| gac cca ttt tcg ctt gcc gat gag ctg tca ctt ctt agt aat aag Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Leu Leu Ser Asn Lys | 345 |
| 85 90 95 | |
| ttg aga gag atg gta ctt gcc gag gtt cca aag ctt gcc tct gct gct Leu Arg Glu Met Val Leu Ala Glu Val Pro Lys Leu Ala Ser Ala Ala | 393 |
| 100 105 110 | |
| gag tac ttc ttc aaa agg ggt gtg caa gga aaa cag ttt cgt tca act Glu Tyr Phe Phe Lys Arg Gly Val Gln Gly Lys Gln Phe Arg Ser Thr | 441 |
| 115 120 125 130 | |
| att ttg ctg ctg atg gcg aca gct ctg gat gta cga gtt cca gaa gca Ile Leu Leu Met Ala Thr Ala Leu Asp Val Arg Val Pro Glu Ala | 489 |
| 135 140 145 | |
| ttg att ggg gaa tca aca gat ata gtc aca tca gaa tta cgc gta agg Leu Ile Gly Glu Ser Thr Asp Ile Val Thr Ser Glu Leu Arg Val Arg | 537 |
| 150 155 160 | |
| caa cgg ggt att gct gaa atc act gaa atg ata cac gtc gca agt cta Gln Arg Gly Ile Ala Glu Ile Thr Glu Met Ile His Val Ala Ser Leu | 585 |
| 165 170 175 | |
| ctg cac gat gat gtc ttg gat gat gcc gat aca agg cgt ggt gtt ggt Leu His Asp Asp Val Leu Asp Asp Ala Asp Thr Arg Arg Gly Val Gly | 633 |
| 180 185 190 | |
| tcc tta aat gtt gta atg ggt aac aag atg tcg gta tta gca gga gac Ser Leu Asn Val Val Met Gly Asn Lys Met Ser Val Leu Ala Gly Asp | 681 |
| 195 200 205 210 | |
| ttc ttg ctc tcc cgg gct tgt ggg gct ctc gct gct tta aag aac aca | 729 |

| | | | |
|---|-----|-----|------|
| Phe Leu Leu Ser Arg Ala Cys Gly Ala Leu Ala Ala Leu Lys Asn Thr | | | |
| 215 | 220 | 225 | |
| gag gtt gta gca tta ctt gca act gct gta gaa cat ctt gtt acc acc ggt | | | 777 |
| Glu Val Val Ala Leu Leu Ala Thr Ala Val Glu His Leu Val Thr Gly | | | |
| 230 | 235 | 240 | |
| gaa acc atg gag ata act agt tca acc gag cag cgt tat agt atg gac | | | 825 |
| Glu Thr Met Glu Ile Thr Ser Ser Thr Glu Gln Arg Tyr Ser Met Asp | | | |
| 245 | 250 | 255 | |
| tac tac atg cag aag aca tat tat aag aca gca tcg cta atc tct aac | | | 873 |
| Tyr Tyr Met Gln Lys Thr Tyr Tyr Lys Thr Ala Ser Leu Ile Ser Asn | | | |
| 260 | 265 | 270 | |
| agc tgc aaa gct gtt gcc gtt ctc act gga caa aca gca gaa gtt gcc | | | 921 |
| Ser Cys Lys Ala Val Ala Val Leu Thr Gly Gln Thr Ala Glu Val Ala | | | |
| 275 | 280 | 285 | 290 |
| gtg tta gct ttt gag tat ggg agg aat ctg ggt tta gca ttc caa tta | | | 969 |
| Val Leu Ala Phe Glu Tyr Gly Arg Asn Leu Gly Leu Ala Phe Gln Leu | | | |
| 295 | 300 | 305 | |
| ata gac gac att ctt gat ttc acg ggc aca tct gcc tct ctc gga aag | | | 1017 |
| Ile Asp Asp Ile Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu Gly Lys | | | |
| 310 | 315 | 320 | |
| gga tcg ttg tca gat att cgc cat gga gtc ata aca gcc cca atc ctc | | | 1065 |
| Gly Ser Leu Ser Asp Ile Arg His Gly Val Ile Thr Ala Pro Ile Leu | | | |
| 325 | 330 | 335 | |
| ttt gcc atg gaa gag ttt cct caa cta cgc gaa gtt gtt gat caa gtt | | | 1113 |
| Phe Ala Met Glu Glu Phe Pro Gln Leu Arg Glu Val Val Asp Gln Val | | | |
| 340 | 345 | 350 | |
| gaa aaa gat cct agg aat gtt gac att gct tta gag tat ctt ggg aag | | | 1161 |
| Glu Lys Asp Pro Arg Asn Val Asp Ile Ala Leu Glu Tyr Leu Gly Lys | | | |
| 355 | 360 | 365 | 370 |
| agc aag gga ata cag agg gca aga gaa tta gcc atg gaa cat gcg aat | | | 1209 |
| Ser Lys Gly Ile Gln Arg Ala Arg Glu Leu Ala Met Glu His Ala Asn | | | |
| 375 | 380 | 385 | |
| cta gca gca gct gca atc ggg tct cta cct gaa aca gac aat gaa gat | | | 1257 |
| Leu Ala Ala Ala Ile Gly Ser Leu Pro Glu Thr Asp Asn Glu Asp | | | |
| 390 | 395 | 400 | |
| gtc aaa aga tcg agg cgg gca ctt att gac ttg acc cat aga gtc atc | | | 1305 |
| Val Lys Arg Ser Arg Arg Ala Leu Ile Asp Leu Thr His Arg Val Ile | | | |
| 405 | 410 | 415 | |
| acc aga aac aag tgagattaag taatgtttct ctctatacac caaaaacattc | | | 1357 |
| Thr Arg Asn Lys | | | |
| 420 | | | |
| ctcatttcat ttgttaggatt ttgttgttcc aattcgtttc acgaa | | | 1402 |

<210> 64

<211> 422

<212> PRT

<213> Arabidopsis thaliana

<400> 64

Met Leu Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg
1 5 10 15

Asn Arg Ser Phe Tyr Gly Ser Ser Gln Ser Leu Ala Ser His Arg Phe
20 25 30

Ala Ile Ile Pro Asp Gln Gly His Ser Cys Ser Asp Ser Pro His Lys
35 40 45

Gly Tyr Val Cys Arg Thr Thr Tyr Ser Leu Lys Ser Pro Val Phe Gly
50 55 60

Gly Phe Ser His Gln Leu Tyr His Gln Ser Ser Ser Leu Val Glu Glu
65 70 75 80

Glu Leu Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Leu Leu Ser
85 90 95

Asn Lys Leu Arg Glu Met Val Leu Ala Glu Val Pro Lys Leu Ala Ser
100 105 110

Ala Ala Glu Tyr Phe Phe Lys Arg Gly Val Gln Gly Lys Gln Phe Arg
115 120 125

Ser Thr Ile Leu Leu Leu Met Ala Thr Ala Leu Asp Val Arg Val Pro
130 135 140

Glu Ala Leu Ile Gly Glu Ser Thr Asp Ile Val Thr Ser Glu Leu Arg
145 150 155 160

Val Arg Gln Arg Gly Ile Ala Glu Ile Thr Glu Met Ile His Val Ala
165 170 175

Ser Leu Leu His Asp Asp Val Leu Asp Asp Ala Asp Thr Arg Arg Gly
180 185 190

Val Gly Ser Leu Asn Val Val Met Gly Asn Lys Met Ser Val Leu Ala
195 200 205

Gly Asp Phe Leu Leu Ser Arg Ala Cys Gly Ala Leu Ala Ala Leu Lys
210 215 220

Asn Thr Glu Val Val Ala Leu Leu Ala Thr Ala Val Glu His Leu Val
225 230 235 240

Thr Gly Glu Thr Met Glu Ile Thr Ser Ser Thr Glu Gln Arg Tyr Ser
245 250 255

Met Asp Tyr Tyr Met Gln Lys Thr Tyr Tyr Lys Thr Ala Ser Leu Ile
260 265 270

Ser Asn Ser Cys Lys Ala Val Ala Val Leu Thr Gly Gln Thr Ala Glu
275 280 285

Val Ala Val Leu Ala Phe Glu Tyr Gly Arg Asn Leu Gly Leu Ala Phe
290 295 300

Gln Leu Ile Asp Asp Ile Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu
305 310 315 320

Gly Lys, Gly Ser Leu Ser Asp Ile Arg His Gly Val Ile Thr Ala Pro
325 330 335

Ile Leu Phe Ala Met Glu Glu Phe Pro Gln Leu Arg Glu Val Val Asp
340 345 350

Gln Val Glu Lys Asp Pro Arg Asn Val Asp Ile Ala Leu Glu Tyr Leu
355 360 365

Gly Lys Ser Lys Gly Ile Gln Arg Ala Arg Glu Leu Ala Met Glu His
370 375 380

Ala Asn Leu Ala Ala Ala Ile Gly Ser Leu Pro Glu Thr Asp Asn
385 390 395 400

Glu Asp Val Lys Arg Ser Arg Arg Ala Leu Ile Asp Leu Thr His Arg
 405 410 415

Val Ile Thr Arg Asn Lys
 420

<210> 65

<211> 1155

<212> DNA

<213> *Arabidopsis thaliana*

<220>

<221> CDS : *[]* *[]*

<222> (1) .. (1155)

<223> *[]* *[]*

| | | |
|---|--|-----|
| <400> 65 | | |
| atg agt gtg agt tgt tgt agg aat ctg ggc aag aca ata aaa aag | | 48 |
| Met Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys | | |
| 1 5 10 15 | | |
| gca ata cct tca cat cat ttg cat ctg aga agt ctt ggt ggg agt ctc | | 96 |
| Ala Ile Pro Ser His His Leu His Ser Leu Arg Ser Leu Gly Gly Ser Leu | | |
| 20 25 30 | | |
| tat cgt cgt cgt atc caa agc tct tca atg gag acc gat ctc aag tca | | 144 |
| Tyr Arg Arg Ile Gln Ser Ser Met Glu Thr Asp Leu Lys Ser | | |
| 35 40 45 | | |
| acc ttt ctc aac gtt tat tct gtt ctc aag tct gac ctt ctt cat gac | | 192 |
| Thr Phe Leu Asn Val Tyr Ser Val Leu Lys Ser Asp Leu Leu His Asp | | |
| 50 55 60 | | |
| cct tcc ttc gaa ttc acc aat gaa tct cgt ctc tgg gtt gat cgg atg | | 240 |
| Pro Ser Phe Glu Phe Thr Asn Glu Ser Arg Leu Trp Val Asp Arg Met | | |
| 65 70 75 80 | | |
| ctg gac tac aat gta cgt gga ggg aaa ctc aat cgg ggt ctc tct gtt | | 288 |
| Leu Asp Tyr Asn Val Arg Gly Gly Lys Leu Asn Arg Gly Leu Ser Val | | |
| 85 90 95 | | |
| gtt gac agt ttc aaa ctt ttg aag caa ggc aat gat ttg act gag caa | | 336 |

120

| | | | | |
|---|-----|-----|-----|------|
| Val Asp Ser Phe Lys Leu Leu Lys Gln Gly Asn Asp Leu Thr Glu Gln | | | | |
| 100 | 105 | 110 | | |
| gag gtt ttc ctc tct tgt gct ctc ggt tgg tgc att gaa tgg ctc caa | | | | 384 |
| Glu Val Phe Leu Ser Cys Ala Leu Gly Trp Cys Ile Glu Trp Leu Gln | | | | |
| 115 | 120 | 125 | | |
| gct tat ttc ctt gtg ctt gat gat att atg gat aac tct gtc act cgc | | | | 432 |
| Ala Tyr Phe Leu Val Leu Asp Asp Ile Met Asp Asn Ser Val Thr Arg | | | | |
| 130 | 135 | 140 | | |
| cgt ggt caa cct tgc tgg ttc aga gtt cct cag gtt ggt atg gtt gcc | | | | 480 |
| Arg Gly Gln Pro Cys Trp Phe Arg Val Pro Gln Val Gly Met Val Ala | | | | |
| 145 | 150 | 155 | 160 | |
| atc aat gat ggg att cta ctt cgc aat cac atc cac agg att ctc aaa | | | | 528 |
| Ile Asn Asp Gly Ile Leu Leu Arg Asn His Ile His Arg Ile Leu Lys | | | | |
| 165 | 170 | 175 | | |
| aag cat ttc cgt gat aag cct tac tat gtt gac ctt gtt gat ttg ttt | | | | 576 |
| Lys His Phe Arg Asp Lys Pro Tyr Tyr Val Asp Leu Val Asp Leu Phe | | | | |
| 180 | 185 | 190 | | |
| aat gag gtt gag ttg caa aca gct tgt ggc cag atg ata gat ttg atc | | | | 624 |
| Asn Glu Val Glu Leu Gln Thr Ala Cys Gly Gln Met Ile Asp Leu Ile | | | | |
| 195 | 200 | 205 | | |
| acc acc ttt gaa gga gaa aag gat ttg gcc aag tac tca ttg tca atc | | | | 672 |
| Thr Thr Phe Glu Gly Glu Lys Asp Leu Ala Lys Tyr Ser Leu Ser Ile | | | | |
| 210 | 215 | 220 | | |
| cac cgt cgt att gtc cag tac aaa acg gct tat tac tca ttt tat ctc | | | | 720 |
| His Arg Arg Ile Val Gln Tyr Lys Thr Ala Tyr Tyr Ser Phe Tyr Leu | | | | |
| 225 | 230 | 235 | 240 | |
| cct gtt gct tgt gcg ttg ctt atg gcg ggc gaa aat ttg gaa aac cat | | | | 768 |
| Pro Val Ala Cys Ala Leu Leu Met Ala Gly Glu Asn Leu Glu Asn His | | | | |
| 245 | 250 | 255 | | |
| att gac gtg aaa aat gtt ctt gtt gac atg gga atc tac ttc caa gtg | | | | 816 |
| Ile Asp Val Lys Asn Val Leu Val Asp Met Gly Ile Tyr Phe Gln Val | | | | |
| 260 | 265 | 270 | | |
| cag gat gat tat ctg gat tgt ttt gct gat ccc gag acg ctt ggc aag | | | | 864 |
| Gln Asp Asp Tyr Leu Asp Cys Phe Ala Asp Pro Glu Thr Leu Gly Lys | | | | |
| 275 | 280 | 285 | | |
| ata gga aca gat ata gaa gat ttc aaa tgc tcg tgg ttg gtg gtt aag | | | | 912 |
| Ile Gly Thr Asp Ile Glu Asp Phe Lys Cys Ser Trp Leu Val Val Lys | | | | |
| 290 | 295 | 300 | | |
| gca tta gag cgc tgc agc gaa gaa caa act aag ata tat gag aac | | | | 960 |
| Ala Leu Glu Arg Cys Ser Glu Glu Gln Thr Lys Ile Leu Tyr Glu Asn | | | | |
| 305 | 310 | 315 | 320 | |
| tat ggt aaa ccc gac cca tcg aac gtt gct aaa gtg aag gat ctc tac | | | | 1008 |

| | | | |
|---|-----|-----|------|
| Tyr Gly Lys Pro Asp Pro Ser Asn Val Ala Lys Val Lys Asp Leu Tyr | | | |
| 325 | 330 | 335 | |
| aaa gag ctg gat ctt gag gga gtt ttc atg gag tat gag agc aaa agc | | | 1056 |
| Lys Glu Leu Asp Leu Glu Gly Val Phe Met Glu Tyr Glu Ser Lys Ser | | | |
| 340 | 345 | 350 | |
| tac gag aag ctg act gga gcg att gag gga cac caa agt aaa gca atc | | | 1104 |
| Tyr Glu Lys Leu Thr Gly Ala Ile Glu Gly His Gln Ser Lys Ala Ile | | | |
| 355 | 360 | 365 | |
| caa gca gtg cta aaa tcc ttc ttg gct aag atc tac aag agg cag aag | | | 1152 |
| Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys | | | |
| 370 | 375 | 380 | |
| tag | | | 1155 |
| <210> 66 | | | |
| <211> 384 | | | |
| <212> PRT | | | |
| <213> <i>Arabidopsis thaliana</i> | | | |
| <400> 66 | | | |
| Met Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys | | | |
| 1 | 5 | 10 | 15 |
| Ala Ile Pro Ser His His Leu His Leu Arg Ser Leu Gly Gly Ser Leu | | | |
| 20 | 25 | 30 | |
| Tyr Arg Arg Arg Ile Gln Ser Ser Ser Met Glu Thr Asp Leu Lys Ser | | | |
| 35 | 40 | 45 | |
| Thr Phe Leu Asn Val Tyr Ser Val Leu Lys Ser Asp Leu Leu His Asp | | | |
| 50 | 55 | 60 | |
| Pro Ser Phe Glu Phe Thr Asn Glu Ser Arg Leu Trp Val Asp Arg Met | | | |
| 65 | 70 | 75 | 80 |
| Leu Asp Tyr Asn Val Arg Gly Gly Lys Leu Asn Arg Gly Leu Ser Val | | | |
| 85 | 90 | 95 | |
| Val Asp Ser Phe Lys Leu Leu Lys Gln Gly Asn Asp Leu Thr Glu Gln | | | |
| 100 | 105 | 110 | |

Glu Val Phe Leu Ser Cys Ala Leu Gly Trp Cys Ile Glu Trp Leu Gln
115 120 125

Ala Tyr Phe Leu Val Leu Asp Asp Ile Met Asp Asn Ser Val Thr Arg
130 135 140

Arg Gly Gln Pro Cys Trp Phe Arg Val Pro Gln Val Gly Met Val Ala
145 150 155 160

Ile Asn Asp Gly Ile Leu Leu Arg Asn His Ile His Arg Ile Leu Lys
165 170 175

Lys His Phe Arg Asp Lys Pro Tyr Tyr Val Asp Leu Val Asp Leu Phe
180 185 190

Asn Glu Val Glu Leu Gln Thr Ala Cys Gly Gln Met Ile Asp Leu Ile
195 200 205 | |

Thr Thr Phe Glu Gly Glu Lys Asp Leu Ala Lys Tyr Ser Leu Ser Ile
210 215 220 | |

His Arg Arg Ile Val Gln Tyr Lys Thr Ala Tyr Tyr Ser Phe Tyr Leu
225 230 235 240

Pro Val Ala Cys Ala Leu Leu Met Ala Gly Glu Asn Leu Glu Asn His
245 250 255

Ile Asp Val Lys Asn Val Leu Val Asp Met Gly Ile Tyr Phe Gln Val
260 265 270

Gln Asp Asp Tyr Leu Asp Cys Phe Ala Asp Pro Glu Thr Leu Gly Lys
275 280 285

Ile Gly Thr Asp Ile Glu Asp Phe Lys Cys Ser Trp Leu Val Val Lys
290 295 300

Ala Leu Glu Arg Cys Ser Glu Glu Gln Thr Lys Ile Leu Tyr Glu Asn
305 310 315 320

Tyr Gly Lys Pro Asp Pro Ser Asn Val Ala Lys Val Lys Asp Leu Tyr
325 330 335

Lys Glu Leu Asp Leu Glu Gly Val Phe Met Glu Tyr Glu Ser Lys Ser
 340 345 350

Tyr Glu Lys Leu Thr Gly Ala Ile Glu Gly His Gln Ser Lys Ala Ile
 355 360 365

Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys
 370 375 380

<210> 67

<211> 1101

<212> DNA

<213> Sinabs alba

<220>

<221> CDS

<222> (1)..(1101)

<223>

| | | |
|--|--|-----|
| <400> 67 | | |
| atg gct tct tca gtg act cct cta ggt tca tgg gtt ctt ctt cac cat | | 48 |
| Met Ala Ser Ser Val Thr Pro Leu Gly Ser Trp Val Leu Leu His His | | |
| 1 5 10 15 | | |
| cat cct tca act atc tta acc caa tcc aga tcc aga tct cct cct tct | | 96 |
| His Pro Ser Thr Ile Leu Thr Gln Ser Arg Ser Arg Ser Pro Pro Ser | | |
| 20 25 30 | | |
| ctc atc acc ctt aaa ccc atc tcc ctc act cca aaa cgc acc gtt tcg | | 144 |
| Leu Ile Thr Leu Lys Pro Ile Ser Leu Thr Pro Lys Arg Thr Val Ser | | |
| 35 40 45 | | |
| tct tct tcc tcc tct tcc ctc atc acc aaa gaa gac aac aac ctc aaa | | 192 |
| Ser Ser Ser Ser Ser Leu Ile Thr Lys Glu Asp Asn Asn Leu Lys | | |
| 50 55 60 | | |
| tcc tct tcc tcc ttc gat ttc atg tct tac atc atc cgc aaa gcc | | 240 |
| Ser Ser Ser Ser Ser Phe Asp Phe Met Ser Tyr Ile Ile Arg Lys Ala | | |
| 65 70 75 80 | | |
| gac tcc gtc aac aaa gcc tta gac tcc gcc gtc cct ctc cgg gag cca | | 288 |

| | | |
|---|-----|-----|
| Asp Ser Val Asn Lys Ala Leu Asp Ser Ala Val Pro Leu Arg Glu Pro | | |
| 85 | 90 | 95 |
| | | |
| ctc aag atc cac gaa gcg atg cgt tac tct ctc ctc gcc gga gga aaa | | 336 |
| Leu Lys Ile His Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys | | |
| 100 | 105 | 110 |
| | | |
| cgc gtc aga cca gtt ctc tgc atc gcc gcg tgc gag cta gtc gga gga | | 384 |
| Arg Val Arg Pro Val Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly | | |
| 115 | 120 | 125 |
| | | |
| gaa gag tct tta gct atg ccg gcg cgt tgc gcc gtg gaa atg atc cac | | 432 |
| Glu Glu Ser Leu Ala Met Pro Ala Arg Cys Ala Val Glu Met Ile His | | |
| 130 | 135 | 140 |
| | | |
| acc atg tcg ttg atc cac gac gac ttg cct tgt atg gat aac gac gat | | 480 |
| Thr Met Ser Leu Ile His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp | | |
| 145 | 150 | 155 |
| 160 | | |
| | | |
| ctc cgc cgc gga aag ccc acg aat cac aaa gtt tac ggc gaa gac gtg | | 528 |
| Leu Arg Arg Gly Lys Pro Thr Asn His Lys Val Tyr Gly Glu Asp Val | | |
| 165 | 170 | 175 |
| | | |
| gcg gtt tta gcc gga gac gcg ctt ctt tcg ttc gcc ttc gag cat tta | | 576 |
| Ala Val Leu Ala Gly Asp Ala Leu Leu Ser Phe Ala Phe Glu His Leu | | |
| 180 | 185 | 190 |
| | | |
| gcg tcg gct acg agc tcg gag gtt tct ccg gcg aga gtg gtt aga gct | | 624 |
| Ala Ser Ala Thr Ser Ser Glu Val Ser Pro Ala Arg Val Val Arg Ala | | |
| 195 | 200 | 205 |
| | | |
| gtg gga gag ttg gct aaa gcc atc ggc acc gaa ggg ctc gtg gcg gga | | 672 |
| Val Gly Glu Leu Ala Lys Ala Ile Gly Thr Glu Gly Leu Val Ala Gly | | |
| 210 | 215 | 220 |
| | | |
| caa gtg gtg gat ata agc agt gaa ggg ttg gac tta aac aac gtc gga | | 720 |
| Gln Val Val Asp Ile Ser Ser Glu Gly Leu Asp Leu Asn Asn Val Gly | | |
| 225 | 230 | 235 |
| 240 | | |
| | | |
| ttg gag cat ttg aag ttt ata cat ttg cat aaa acg gcg gcg ttg ctt | | 768 |
| Leu Glu His Leu Lys Phe Ile His Leu His Lys Thr Ala Ala Leu Leu | | |
| 245 | 250 | 255 |
| | | |
| gaa gct tca gcg gtt ttg ggt ggg atc atc ggt gga ggg agt gat gaa | | 816 |
| Glu Ala Ser Ala Val Leu Gly Gly Ile Ile Gly Gly Ser Asp Glu | | |
| 260 | 265 | 270 |
| | | |
| gag atc gag agg ctg agg aag ttc gcg agg tgt att ggg ttg ttg ttt | | 864 |
| Glu Ile Glu Arg Leu Arg Lys Phe Ala Arg Cys Ile Gly Leu Leu Phe | | |
| 275 | 280 | 285 |
| | | |
| cag gtg gtt gat gat atc ttg gac gtg acg aaa tcg tct caa gaa ctg | | 912 |
| Gln Val Val Asp Asp Ile Leu Asp Val Thr Lys Ser Ser Gln Glu Leu | | |
| 290 | 295 | 300 |
| | | |
| ggg aaa acc gct ggg aaa gat ttg att gct gat aag ttg act tat ccg | | 960 |

125

Gly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp Lys Leu Thr Tyr Pro
305 310 315 320

aag ctc atg ggt ttg gag aaa tcg aga gag ttc gct gag aag ttg aat 1008
Lys Leu Met Gly Leu Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn
325 330 335

aca gag gca cgt gat cag ctt tta ggg ttt gat tcc gac aag gtt gct 1056
Thr Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp Ser Asp Lys Val Ala
340 345 350

cct ttg ttg gct ttg gct aat tac att gcc aat aga cag aac tga 1101
Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Asn Arg Gln Asn
355 360 365

<210> 68

<211> 366

<212> PRT

<213> Sinabs alba

<400> 68

Met Ala Ser Ser Val Thr Pro Leu Gly Ser Trp Val Leu Leu His His
1 5 10 15

His Pro Ser Thr Ile Leu Thr Gln Ser Arg Ser Arg Ser Pro Pro Ser
20 25 30

Leu Ile Thr Leu Lys Pro Ile Ser Leu Thr Pro Lys Arg Thr Val Ser
35 40 45

Ser Ser Ser Ser Ser Leu Ile Thr Lys Glu Asp Asn Asn Leu Lys
50 55 60

Ser Ser Ser Ser Phe Asp Phe Met Ser Tyr Ile Ile Arg Lys Ala
65 70 75 80

Asp Ser Val Asn Lys Ala Leu Asp Ser Ala Val Pro Leu Arg Glu Pro
85 90 95

Leu Lys Ile His Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys
100 105 110

Arg Val Arg Pro Val Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly
115 120 125

Glu Glu Ser Leu Ala Met Pro Ala Arg Cys Ala Val Glu Met Ile His
130 135 140

Thr Met Ser Leu Ile His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp
145 150 155 160

Leu Arg Arg Gly Lys Pro Thr Asn His Lys Val Tyr Gly Glu Asp Val.
165 170 175

Ala Val Leu Ala Gly Asp Ala Leu Leu Ser Phe Ala Phe Glu His Leu
180 185 190

Ala Ser Ala Thr Ser Ser Glu Val Ser Pro Ala Arg Val Val Arg Ala
195 200 205

Val Gly Glu Leu Ala Lys Ala Ile Gly Thr Glu Gly Leu Val Ala Gly
210 215 220

Gln Val Val Asp Ile Ser Ser Glu Gly Leu Asp Leu Asn Asn Val Gly
225 230 235 240

Leu Glu His Leu Lys Phe Ile His Leu His Lys Thr Ala Ala Leu Leu
245 250 255

Glu Ala Ser Ala Val Leu Gly Gly Ile Ile Gly Gly Ser Asp Glu
260 265 270

Glu Ile Glu Arg Leu Arg Lys Phe Ala Arg Cys Ile Gly Leu Leu Phe
275 280 285

Gln Val Val Asp Asp Ile Leu Asp Val Thr Lys Ser Ser Gln Glu Leu
290 295 300

Gly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp Lys Leu Thr Tyr Pro
305 310 315 320

Lys Leu Met Gly Leu Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn
325 330 335

Thr Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp Ser Asp Lys Val Ala
 340 345 350

Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Asn Arg Gln Asn
 355 360 365

<210> 69

<211> 930

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)...(930)

<223>

| | | | |
|--|--|--|-----|
| <400> 69 | | | |
| atg aat aat ccg tcg tta ctc aat cat gcg gtc gaa acg atg gca gtt | | | 48 |
| Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val | | | |
| 1 5 10 15 | | | |
| ggc tcg aaa agt ttt gcg aca gcc tca aag tta ttt gat gca aaa acc | | | 96 |
| Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr | | | |
| 20 25 30 | | | |
| cgg cgc agc gta ctg atg ctc tac gcc tgg tgc cgc cat tgt gac gat | | | 144 |
| Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp | | | |
| 35 40 45 | | | |
| gtt att gac gat cag acg ctg ggc ttt cag gcc cgg cag cct gcc tta | | | 192 |
| Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu | | | |
| 50 55 60 | | | |
| caa acg ccc gaa caa cgt ctg atg caa ctt gag atg aaa acg cgc cag | | | 240 |
| Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln | | | |
| 65 70 75 80 | | | |
| gcc tat gca gga tcg cag atg cac gaa ccg gcg ttt gcg gct ttt cag | | | 288 |
| Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln | | | |
| 85 90 95 | | | |
| gaa gtg gct atg gct cat gat atc gcc ccg gct tac gcg ttt gat cat | | | 336 |
| Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His | | | |
| 100 105 110 | | | |

| | |
|---|-----|
| ctg gaa ggc ttc gcc atg gat gta cgc gaa gcg caa tac agc caa ctg Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu 115 120 125 | 384 |
| gat gat acg ctg cgc tat tgc tat cac gtt gca ggc gtt gtc ggc ttg Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu 130 135 140 | 432 |
| atg atg gcg caa atc atg ggc gtg cgg gat aac gcc acg ctg gac cgc Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg 145 150 155 160 | 480 |
| gcc tgt gac ctt ggg ctg gca ttt cag ttg acc aat att gct cgc gat Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp 165 170 175 | 528 |
| att gtg gac gat gcg cat gcg ggc cgc tgt tat ctg ccc gca agc tgg Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp 180 185 190 | 576 |
| ctg gag cat gaa ggt ctg aac aaa gag aat tat gcg gca cct gaa aac Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn 195 200 205 | 624 |
| cgt cag gcg ctg agc cgt atc gcc cgt cgt ttg gtg cag gaa gca gaa Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val Gln Glu Ala Glu 210 215 220 | 672 |
| cct tac tat ttg tct gcc aca gcc ggc ctg gca ggg ttg ccc ctg cgt Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu Pro Leu Arg 225 230 235 240 | 720 |
| tcc gcc tgg gca atc gct acg gcg aag cag gtt tac cgg aaa ata ggt Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg Lys Ile Gly 245 250 255 | 768 |
| gtc aaa gtt gaa cag gcc ggt cag caa gcc tgg gat cag cgg cag tca Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln Arg Gln Ser 260 265 270 | 816 |
| acg acc acg ccc gaa aaa tta acg ctg ctg ctg gcc gcc tct ggt cag Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Ala Ala Ser Gly Gln 275 280 285 | 864 |
| gcc ctt act tcc cgg atg cgg gct cat cct ccc cgc cct gcg cat ctc Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro Ala His Leu 290 295 300 | 912 |
| tgg cag cgc ccg ctc tag Trp Gln Arg Pro Leu 305 | 930 |

<211> 309

<212> PRT

<213> Erwinia uredovora

<400> 70

Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val
1 5 10 15

Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr
20 25 30

Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp
35 40 45

Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu
50 55 60

Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln
65 70 75 80

Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln
85 90 95

Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His
100 105 110

Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu
115 120 125

Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu
130 135 140

Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg
145 150 155 160

Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp
165 170 175

Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp
180 185 190

Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn
195 200 205

Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val Gln Glu Ala Glu
210 215 220

Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu Pro Leu Arg
225 230 235 240

Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg Lys Ile Gly
245 250 255

Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln Arg Gln Ser
260 265 270

Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Leu Ala Ala Ser Gly Gln
275 280 285

Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro Ala His Leu
290 295 300

Trp Gln Arg Pro Leu
305

<210> 71

<211> 1479

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(1479)

<223>

<400> 71

atg aaa cca act acg gta att ggt gca ggc ttc ggt ggc ctg gca ctg

| | | | | | |
|---|---|----|----|--|-----|
| Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu | | | | | |
| 1 | 5 | 10 | 15 | | |
| gca att cgt cta caa gct gcg ggg atc ccc gtc tta ctg ctt gaa caa | | | | | 96 |
| Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Glu Gln | | | | | |
| 20 25 30 | | | | | |
| cgt gat aaa ccc ggc ggt cgg gct tat gtc tac gag gat cag ggg ttt | | | | | 144 |
| Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Glu Asp Gln Gly Phe | | | | | |
| 35 40 45 | | | | | |
| acc ttt gat gca ggc ccg acg gtt atc acc gat ccc agt gcc att gaa | | | | | 192 |
| Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu | | | | | |
| 50 55 60 | | | | | |
| gaa ctg ttt gca ctg gca gga aaa cag tta aaa gag tat gtc gaa ctg | | | | | 240 |
| Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Glu Tyr Val Glu Leu | | | | | |
| 65 70 75 80 | | | | | |
| ctg ccg gtt acg ccg ttt tac cgc ctg tgt tgg gag tca ggg aag gtc | | | | | 288 |
| Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val | | | | | |
| 85 90 95 | | | | | |
| ttt aat tac gat aac gat caa acc cgg ctc gaa gcg cag att cag cag | | | | | 336 |
| Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu Ala Gln Ile Gln Gln | | | | | |
| 100 105 110 | | | | | |
| ttt aat ccc cgc gat gtc gaa ggt tat cgt cag ttt ctg gac tat tca | | | | | 384 |
| Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln Phe Leu Asp Tyr Ser | | | | | |
| 115 120 125 | | | | | |
| cgc gcg gtg ttt aaa gaa ggc tat cta aag ctc ggt act gtc cct ttt | | | | | 432 |
| Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe | | | | | |
| 130 135 140 | | | | | |
| tta tcg ttc aga gac atg ctt cgc gcc gca cct caa ctg gcg aaa ctg | | | | | 480 |
| Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu | | | | | |
| 145 150 155 160 | | | | | |
| cag gca tgg aga agc gtt tac agt aag gtt gcc agt tac atc gaa gat | | | | | 528 |
| Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Ser Tyr Ile Glu Asp | | | | | |
| 165 170 175 | | | | | |
| gaa cat ctg cgc cag gcg ttt tct ttc cac tcg ctg ttg gtg ggc ggc | | | | | 576 |
| Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly | | | | | |
| 180 185 190 | | | | | |
| aat ccc ttc gcc acc tca tcc att tat acg ttg ata cac gcg ctg gag | | | | | 624 |
| Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu | | | | | |
| 195 200 205 | | | | | |
| cgt gag tgg ggc gtc tgg ttt ccg cgt ggc ggc acc ggc gca tta gtt | | | | | 672 |
| Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val | | | | | |
| 210 215 220 | | | | | |
| cag ggg atg ata aag ctg ttt cag gat ctg ggt ggc gaa gtc gtg tta | | | | | 720 |
| | | | | | |

| | | | |
|---|-----|-----|------|
| Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu | | | |
| 225 | 230 | 235 | 240 |
| aac gcc aga gtc agc cat atg gaa acg aca gga aac aag att gaa gcc | | | 768 |
| Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala | | | |
| 245 | 250 | 255 | |
| gtg cat tta gag gac ggt cgc agg ttc ctg acg caa gcc gtc gcg tca | | | 816 |
| Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser | | | |
| 260 | 265 | 270 | |
| aat gca gat gtg gtt cat acc tat cgc gac ctg tta agc cag cac cct | | | 864 |
| Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro | | | |
| 275 | 280 | 285 | |
| gcc gcg gtt aag cag tcc aac aaa ctg cag act aag cgc atg agt aac | | | 912 |
| Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn | | | |
| 290 | 295 | 300 | |
| tct ctg ttt gtg ctc tat ttt ggt ttg aat cac cat cat gat cag ctc | | | 960 |
| Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu | | | |
| 305 | 310 | 315 | 320 |
| gcg cat cac acg gtt tgt ttc ggc ccg cgt tac cgc gag ctg att gac | | | 1008 |
| Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp | | | |
| 325 | 330 | 335 | |
| gaa att ttt aat cat gat ggc ctc gca gag gac ttc tca ctt tat ctg | | | 1056 |
| Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu | | | |
| 340 | 345 | 350 | |
| cac gcg ccc tgt gtc acg gat tcg tca ctg gcg cct gaa ggt tgc ggc | | | 1104 |
| His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly | | | |
| 355 | 360 | 365 | |
| agt tac tat gtg ttg gcg ccg gtg ccg cat tta ggc acc gcg aac ctc | | | 1152 |
| Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu | | | |
| 370 | 375 | 380 | |
| gac tgg acg gtt gag ggg cca aaa cta cgc gac cgt att ttt gcg tac | | | 1200 |
| Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr | | | |
| 385 | 390 | 395 | 400 |
| ctt gag cag cat tac atg cct ggc tta cgg agt cag ctg gtc acg cac | | | 1248 |
| Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His | | | |
| 405 | 410 | 415 | |
| cgg atg ttt acg ccg ttt gat ttt cgc gac cag ctt aat gcc tat cat | | | 1296 |
| Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His | | | |
| 420 | 425 | 430 | |
| ggc tca gcc ttt tct gtg gag ccc gtt ctt acc cag agc gcc tgg ttt | | | 1344 |
| Gly Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe | | | |
| 435 | 440 | 445 | |
| cgg ccg cat aac cgc gat aaa acc att act aat ctc tac ctg gtc ggc | | | 1392 |

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly
450 455 460 1440

gca ggc acg cat ccc ggc gca ggc att cct ggc gtc atc ggc tcg gca
Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala
465 470 475 480 1440

aaa gcg aca gca ggt ttg atg ctg gag gat ctg ata tga
Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile
485 490 1479

<210> 72

<211> 492

<212> PRT

<213> Erwinia uredovora

<400> 72

Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu
1 5 10 15

Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Glu Gln
20 25 30

Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Glu Asp Gln Gly Phe
35 40 45

Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu
50 55 60

Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Glu Tyr Val Glu Leu
65 70 75 80

Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val
85 90 95

Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu Ala Gln Ile Gln Gln
100 105 110

Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln Phe Leu Asp Tyr Ser
115 120 125

134

Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe
130 135 140

Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu
145 150 155 160

Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Ser Tyr Ile Glu Asp
165 170 175

Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly
180 185 190

Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu
195 200 205

Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val
210 215 220

Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu
225 230 235 240

Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala
245 250 255

Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser
260 265 270

Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro
275 280 285

Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn
290 295 300

Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu
305 310 315 320

Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp
325 330 335

Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu
340 345 350

135

His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly
355 360 365

Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu
370 375 380

Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr
385 390 395 400

Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His
405 410 415

Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His
420 425 430

Gly Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe
435 440 445

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly
450 455 460

[]
Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala
465 470 475 480

Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile
485 490

<210> 73

<211> 1725

<212> DNA

<213> Narcissus pseudonarcissus

<220>

<221> CDS

<222> (1)...(1725)

<223>

<400> 73
atg gct tct tcc act tgt tta att cat tct tcc tct ttt ggg gtt gga 48
Met Ala Ser Ser Thr Cys Leu Ile His Ser Ser Ser Phe Gly Val Gly
1 5 10 15

gga aag aaa gtg aag atg aac acg atg att cga tcg aag ttg ttt tca 96
Gly Lys Lys Val Lys Met Asn Thr Met Ile Arg Ser Lys Leu Phe Ser
20 25 30

att cgg tcg gct ttg gac act aag gtg tct gat atg agc gtc aat gct 144.
Ile Arg Ser Ala Leu Asp Thr Lys Val Ser Asp Met Ser Val Asn Ala
35 40 45

cca aaa gga ttg ttt cca cca gag cct gag cac tac agg ggg cca aag 192
Pro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro Lys
50 55 60

ctt aaa gtg gct atc att gga gct ggg ctc gct ggc atg tca act gca 240
Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr Ala
65 70 75 80

gtg gag ctt ttg gat caa ggg cat gag gtt gac ata tat gaa tcc aga 288
Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser Arg
85 90 95

caa ttt att ggt ggt aaa gtc ggt tct ttt gta gat aag cgt gga aac 336
Gln Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Lys Arg Gly Asn
100 105 110

cat att gaa atg gga ctc cat gtg ttt ttt ggt tgc tat aac aat ctt 384
His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn Leu
115 120 125

ttc aga ctt atg aaa aag gta ggt gca gat gaa aat tta ctg gtg aag 432
Phe Arg Leu Met Lys Lys Val Gly Ala Asp Glu Asn Leu Leu Val Lys
130 135 140

gat cat act cat acc ttt gta aac cga ggt gga gaa att ggt gaa ctt 480
Asp His Thr His Thr Phe Val Asn Arg Gly Gly Glu Ile Gly Glu Leu
145 150 155 160

gat ttc cga ctt ccg atg ggt gca cca tta cat ggt att cgt gca ttt 528
Asp Phe Arg Leu Pro Met Gly Ala Pro Leu His Gly Ile Arg Ala Phe
165 170 175

cta aca act aat caa ctg aag cct tat gat aaa gca agg aat gct gtg 576
Leu Thr Thr Asn Gln Leu Lys Pro Tyr Asp Lys Ala Arg Asn Ala Val
180 185 190

gct ctt gcc ctt agc cca gtt gta cgt gct ctt att gat cca aat ggt 624
Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Ile Asp Pro Asn Gly
195 200 205

gca atg cag gat ata agg aac tta gat aat att agc ttt tct gat tgg 672
Ala Met Gln Asp Ile Arg Asn Leu Asp Asn Ile Ser Phe Ser Asp Trp
210 215 220

| | |
|---|------|
| ttc tta tcc aaa ggc ggt acc cgc atg agc atc caa agg atg tgg gat Phe Leu Ser Lys Gly Gly Thr Arg Met Ser Ile Gln Arg Met Trp Asp 225 230 235 240 | 720 |
| cca gtt gct tat gcc ctc gga ttt att gac tgt gat aat atc agt gcc Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser Ala 245 250 255 | 768 |
| cgt tgt atg ctt act ata ttt tct cta ttt gct act aag aca gaa gct Arg Cys Met Leu Thr Ile Phe Ser Leu Phe Ala Thr Lys Thr Glu Ala 260 265 270 | 816 |
| tct ctg ttg cgt atg ttg aag ggt tcg cct gat gtt tac tta agc ggt Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser Gly 275 280 285 | 864 |
| cct ata aga aag tat att aca gat aaa ggt gga agg ttt cac cta agg Pro Ile Arg Lys Tyr Ile Thr Asp Lys Gly Arg Phe His Leu Arg 290 295 300 | 912 |
| tgg ggg tgt aga gag ata ctt tat gat gaa cta tca aat ggc gac aca Trp Gly Cys Arg Glu Ile Leu Tyr Asp Glu Leu Ser Asn Gly Asp Thr 305 310 315 320 | 960 |
| tat atc aca ggc att gca atg tcg aag gct acc aat aaa aaa ctt gtg Tyr Ile Thr Gly Ile Ala Met Ser Lys Ala Thr Asn Lys Lys Leu Val 325 330 335 | 1008 |
| aaa gct gac gtg tat gtt gca gca tgt gat gtt cct gga ata aaa agg Lys Ala Asp Val Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys Arg 340 345 350 | 1056 |
| ttg atc cca tcg gag tgg aga gaa tgg gat cta ttt gac aat atc tat Leu Ile Pro Ser Glu Trp Arg Glu Trp Asp Leu Phe Asp Asn Ile Tyr 355 360 365 | 1104 |
| aaa cta gtt gga gtt cca gtt gtc act gtt cag ctt agg tac aat ggt Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn Gly 370 375 380 | 1152 |
| tgg gtg aca gag atg caa gat ctg gaa aaa tca agg cag ttg aga gct Trp Val Thr Glu Met Gln Asp Leu Glu Lys Ser Arg Gln Leu Arg Ala 385 390 395 400 | 1200 |
| gca gta gga ttg gat aat ctt ctt tat act cca gat gca gac ttt tct Ala Val Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe Ser 405 410 415 | 1248 |
| tgt ttt tct gat ctt gca ctc tcg tcg cct gaa gat tat tat att gaa Cys Phe Ser Asp Leu Ala Leu Ser Ser Pro Glu Asp Tyr Tyr Ile Glu 420 425 430 | 1296 |
| gga caa ggg tcc cta ata cag gct ctc acg cca ggg gat cca tac Gly Gln Gly Ser Leu Ile Gln Ala Val Leu Thr Pro Gly Asp Pro Tyr 435 440 445 | 1344 |

| | |
|--|------|
| atg ccc cta cct aat gat gca att ata gaa aga gtt cggtt Met Pro Leu Pro Asn Asp Ala Ile Ile Glu Arg Val Arg Lys Gln Val | 1392 |
| 450 455 460 | |
| ttg gat tta ttc cca tcc tct caa ggc ctg gaa gtt cta tgg tct tcg Leu Asp Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Leu Trp Ser Ser | 1440 |
| 465 470 475 480 | |
| gtg gtt aaa atc gga caa tcc cta tat cgg gag ggg cct gga aag gac Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys Asp | 1488 |
| 485 490 495 | |
| cca ttc aga cct gat cag aag aca cca gta aaa aat ttc ttc ctt gca Pro Phe Arg Pro Asp Gln Lys Thr Pro Val Lys Asn Phe Phe Leu Ala | 1536 |
| 500 505 510 | |
| ggc tca tac acc aaa cag gat tac att gac agt atg gaa gga gcg acc Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala Thr | 1584 |
| 515 520 525 | |
| cta tcg ggg aga caa gca gct gca tat atc tgc agc gcc ggt gaa gat Leu Ser Gly Arg Gln Ala Ala Ala Tyr Ile Cys Ser Ala Gly Glu Asp | 1632 |
| 530 535 540 | |
| ctg gca gca ctt cgc aag aag atc gct gct gat cat cca gag caa ctg Leu Ala Ala Leu Arg Lys Lys Ile Ala Ala Asp His Pro Glu Gln Leu | 1680 |
| 545 550 555 560 | |
| atc aac aaa gat tct aac gtg tcg gat gaa ctg agt ctc gta taa Ile Asn Lys Asp Ser Asn Val Ser Asp Glu Leu Ser Leu Val | 1725 |
| 565 570 | |
| <210> 74 | |
| <211> 574 | |
| <212> PRT | |
| <213> Narcissus pseudonarcissus | |
| <400> 74 | |
| Met Ala Ser Ser Thr Cys Leu Ile His Ser Ser Ser Phe Gly Val Gly 1 5 10 15 | |
| Gly Lys Lys Val Lys Met Asn Thr Met Ile Arg Ser Lys Leu Phe Ser 20 25 30 | |
| Ile Arg Ser Ala Leu Asp Thr Lys Val Ser Asp Met Ser Val Asn Ala 35 40 45 | |

Pro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro Lys
50 55 60

Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr Ala
65 70 75 80

Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser Arg
85 90 95

Gln Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Lys Arg Gly Asn
100 105 110

His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn Leu
115 120 125

Phe Arg Leu Met Lys Lys Val Gly Ala Asp Glu Asn Leu Leu Val Lys
130 135 140

Asp His Thr His Thr Phe Val Asn Arg Gly Gly Glu Ile Gly Glu Leu
145 150 155 160

Asp Phe Arg Leu Pro Met Gly Ala Pro Leu His Gly Ile Arg Ala Phe
165 170 175

Leu Thr Thr Asn Gln Leu Lys Pro Tyr Asp Lys Ala Arg Asn Ala Val
180 185 190

Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Ile Asp Pro Asn Gly
195 200 205

Ala Met Gln Asp Ile Arg Asn Leu Asp Asn Ile Ser Phe Ser Asp Trp
210 215 220

Phe Leu Ser Lys Gly Gly Thr Arg Met Ser Ile Gln Arg Met Trp Asp
225 230 235 240

Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser Ala
245 250 255

Arg Cys Met Leu Thr Ile Phe Ser Leu Phe Ala Thr Lys Thr Glu Ala
260 265 270

Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser Gly
275 280 285

Pro Ile Arg Lys Tyr Ile Thr Asp Lys Gly Gly Arg Phe His Leu Arg
290 295 300

Trp Gly Cys Arg Glu Ile Leu Tyr Asp Glu Leu Ser Asn Gly Asp Thr
305 310 315 320

Tyr Ile Thr Gly Ile Ala Met Ser Lys Ala Thr Asn Lys Lys Leu Val
325 330 335

Lys Ala Asp Val Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys Arg
340 345 350

Leu Ile Pro Ser Glu Trp Arg Glu Trp Asp Leu Phe Asp Asn Ile Tyr
355 360 365

Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn Gly
370 375 380

Trp Val Thr Glu Met Gln Asp Leu Glu Lys Ser Arg Gln Leu Arg Ala
385 390 395 400

Ala Val Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe Ser
405 410 415

Cys Phe Ser Asp Leu Ala Leu Ser Ser Pro Glu Asp Tyr Tyr Ile Glu
420 425 430

Gly Gln Gly Ser Leu Ile Gln Ala Val Leu Thr Pro Gly Asp Pro Tyr
435 440 445

Met Pro Leu Pro Asn Asp Ala Ile Ile Glu Arg Val Arg Lys Gln Val
450 455 460

Leu Asp Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Leu Trp Ser Ser
465 470 475 480

Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys Asp
485 490 495

Pro Phe Arg Pro Asp Gln Lys Thr Pro Val Lys Asn Phe Phe Leu Ala
500 505 510

Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala Thr
515 520 525

Leu Ser Gly Arg Gln Ala Ala Ala Tyr Ile Cys Ser Ala Gly Glu Asp
530 535 540

Leu Ala Ala Leu Arg Lys Lys Ile Ala Ala Asp His Pro Glu Gln Leu
545 550 555 560

Ile Asn Lys Asp Ser Asn Val Ser Asp Glu Leu Ser Leu Val
565 570

<210> 75

<211> 1848

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(1848)

<223>

atg tgt acc ttg agt ttt atg tat cct aat tca ctt ctt gat ggt acc 48
Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr
1 5 10 15

tgc aag act gta gct ttg ggt gat agc aaa cca aga tac aat aaa cag 96
Cys Lys Thr Val Ala Leu Gly Asp Ser Lys Pro Arg Tyr Asn Lys Gln
20 25 30

aga agt tct tgt ttt gac cct ttg ata att gga aat tgt act gat cag 144
Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln
35 40 45

cag cag ctt tgt ggc ttg agt tgg ggg gtg gac aag gct aag gga aga 192

| | | |
|---|-----|-----|
| Gln Gln Leu Cys Gly Leu Ser Trp Gly Val Asp Lys Ala Lys Gly Arg | | |
| 50 | 55 | 60 |
| | | |
| aga ggg ggt act gtt tcc aat ttg aaa gca gtt gta gat gta gac aaa | | 240 |
| Arg Gly Gly Thr Val Ser Asn Leu Lys Ala Val Val Asp Val Asp Lys | | |
| 65 | 70 | 75 |
| | | |
| aga gtg gag agc tat ggc agt agt gat gta gaa gga aat gag agt ggc | | 288 |
| Arg Val Glu Ser Tyr Gly Ser Ser Asp Val Glu Gly Asn Glu Ser Gly | | |
| 85 | 90 | 95 |
| | | |
| agc tat gat gcc att gtt ata ggt tca gga ata ggt gga ttg gtg gca | | 336 |
| Ser Tyr Asp Ala Ile Val Ile Gly Ser Gly Ile Gly Gly Leu Val Ala | | |
| 100 | 105 | 110 |
| | | |
| gcg acg cag ctg gcg gtt aag gga gct aag gtt tta gtt ctg gag aag | | 384 |
| Ala Thr Gln Leu Ala Val Lys Gly Ala Lys Val Leu Val Leu Glu Lys | | |
| 115 | 120 | 125 |
| | | |
| tat gtt att cct ggt gga agc tct ggc ttt tac gag agg gat ggt tat | | 432 |
| Tyr Val Ile Pro Gly Gly Ser Ser Gly Phe Tyr Glu Arg Asp Gly Tyr | | |
| 130 | 135 | 140 |
| | | |
| aag ttt gat gtt ggt tca tca gtg atg ttt gga ttc agt gat aag gga | | 480 |
| Lys Phe Asp Val Gly Ser Ser Val Met Phe Gly Phe Ser Asp Lys Gly | | |
| 145 | 150 | 155 |
| | | |
| 160 | | |
| aac ctc aat tta att act caa gca ttg gca gca gta gga cgt aaa tta | | 528 |
| Asn Leu Asn Leu Ile Thr Gln Ala Leu Ala Val Gly Arg Lys Leu | | |
| 165 | 170 | 175 |
| | | |
| gaa gtt ata cct gac cca aca act gta cat ttc cac ctg cca aat gac | | 576 |
| Glu Val Ile Pro Asp Pro Thr Thr Val His Phe His Leu Pro Asn Asp | | |
| 180 | 185 | 190 |
| | | |
| ctt tct gtt cgt ata cac cga gag tat gat gac ttc att gaa gag ctt | | 624 |
| Leu Ser Val Arg Ile His Arg Glu Tyr Asp Asp Phe Ile Glu Glu Leu | | |
| 195 | 200 | 205 |
| | | |
| gtg agt aaa ttt cca cat gaa aag gaa ggg att atc aaa ttt tac agt | | 672 |
| Val Ser Lys Phe Pro His Glu Lys Glu Gly Ile Ile Lys Phe Tyr Ser | | |
| 210 | 215 | 220 |
| | | |
| gaa tgc tgg aag atc ttt aat tct ctg aat tca ttg gaa ctg aag tct | | 720 |
| Glu Cys Trp Lys Ile Phe Asn Ser Leu Asn Ser Leu Glu Leu Lys Ser | | |
| 225 | 230 | 235 |
| | | |
| 240 | | |
| ttg gag gaa ccc atc tac ctt ttt ggc cag ttc ttt aag aag ccc ctt | | 768 |
| Leu Glu Glu Pro Ile Tyr Leu Phe Gly Gln Phe Phe Lys Lys Pro Leu | | |
| 245 | 250 | 255 |
| | | |
| gaa tgc ttg act ctt gcc tac tat ttg ccc cag aat gct ggt agc atc | | 816 |
| Glu Cys Leu Thr Leu Ala Tyr Tyr Leu Pro Gln Asn Ala Gly Ser Ile | | |
| 260 | 265 | 270 |
| | | |
| gct cgg aag tat ata aga gat cct ggg ttg ctg tct ttt ata gat gca | | 864 |

| | | | | |
|---|-----|-----|-----|------|
| Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala | | | | |
| 275 | 280 | 285 | | |
| gag tgc ttt atc gtg agt aca gtt aat gca tta caa aca cca atg atc | | | | 912 |
| Glu Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gln Thr Pro Met Ile | | | | |
| 290 | 295 | 300 | | |
| aat gca agc atg gtt cta tgt gac aga cat ttt ggc gga atc aac tac | | | | 960 |
| Asn Ala Ser Met Val Leu Cys Asp Arg His Phe Gly Gly Ile Asn Tyr | | | | |
| 305 | 310 | 315 | 320 | |
| ccc gtt ggt gga gtt ggc gag atc gcc aaa tcc tta gca aaa ggc ttg | | | | 1008 |
| Pro Val Gly Gly Val Gly Glu Ile Ala Lys Ser Leu Ala Lys Gly Leu | | | | |
| 325 | 330 | 335 | | |
| gat gat cac gga agt cag ata ctt tat agg gca aat gtt aca agt atc | | | | 1056 |
| Asp Asp His Gly Ser Gln Ile Leu Tyr Arg Ala Asn Val Thr Ser Ile | | | | |
| 340 | 345 | 350 | | |
| att ttg gac aat ggc aaa gct gtg gga gtg aag ctt tct gac ggg agg | | | | 1104 |
| Ile Leu Asp Asn Gly Lys Ala Val Gly Val Lys Leu Ser Asp Gly Arg | | | | |
| 355 | 360 | 365 | | |
| aag ttt tat gct aaa acc ata gta tcg aat gct acc aga tgg gat act | | | | 1152 |
| Lys Phe Tyr Ala Lys Thr Ile Val Ser Asn Ala Thr Arg Trp Asp Thr | | | | |
| 370 | 375 | 380 | | |
| ttt gga aag ctt tta aaa gct gag aat ctg cca aaa gaa gaa gaa aat | | | | 1200 |
| Phe Gly Lys Leu Leu Lys Ala Glu Asn Leu Pro Lys Glu Glu Glu Asn | | | | |
| 385 | 390 | 395 | 400 | |
| ttc cag aaa gct tat gta aaa gca cct tct ttt ctt tct att cat atg | | | | 1248 |
| Phe Gln Lys Ala Tyr Val Lys Ala Pro Ser Phe Leu Ser Ile His Met | | | | |
| 405 | 410 | 415 | | |
| gga gtt aaa gca gat gta ctc cca cca gac aca gat tgt cac cat ttt | | | | 1296 |
| Gly Val Lys Ala Asp Val Leu Pro Pro Asp Thr Asp Cys His His Phe | | | | |
| 420 | 425 | 430 | | |
| gtc ctc gag gat tat tgg aca aat ttg gag aaa cca tat gga agt ata | | | | 1344 |
| Val Leu Glu Asp Asp Trp Thr Asn Leu Glu Lys Pro Tyr Gly Ser Ile | | | | |
| 435 | 440 | 445 | | |
| ttc ttg agt att cca aca gtt ctt gat tcc tca ttg gcc cca gaa gga | | | | 1392 |
| Phe Leu Ser Ile Pro Thr Val Leu Asp Ser Ser Leu Ala Pro Glu Gly | | | | |
| 450 | 455 | 460 | | |
| cac cat att ctt cac att ttt aca aca tcg agc att gaa gat tgg gag | | | | 1440 |
| His His Ile Leu His Ile Phe Thr Thr Ser Ser Ile Glu Asp Trp Glu | | | | |
| 465 | 470 | 475 | 480 | |
| gga ctc tct ccg aaa gac tat gaa gcg aag aaa gag gtt gtt gct gaa | | | | 1488 |
| Gly Leu Ser Pro Lys Asp Tyr Glu Ala Lys Lys Glu Val Val Ala Glu | | | | |
| 485 | 490 | 495 | | |
| agg att ata agc aga ctt gaa aaa aca ctc ttc cca ggg ctt aag tca | | | | 1536 |

| | | | | |
|---|-----|-----|-----|------|
| Arg Ile Ile Ser Arg Leu Glu Lys Thr Leu Phe Pro Gly Leu Lys Ser | | | | |
| 500 | 505 | 510 | | |
| tct att ctc ttt aag gag gtg gga act cca aag acc cac aga cga tac | | | | 1584 |
| Ser Ile Leu Phe Lys Glu Val Gly Thr Pro Lys Thr His Arg Arg Tyr | | | | |
| 515 | 520 | 525 | | |
| ctt gct cgt gat agt ggt acc tat gga cca atg cca cgc gga aca cct | | | | 1632 |
| Leu Ala Arg Asp Ser Gly Thr Tyr Gly Pro Met Pro Arg Gly Thr Pro | | | | |
| 530 | 535 | 540 | | |
| aag gga ctc ctg gga atg cct ttc aat acc act gct ata gat ggt cta | | | | 1680 |
| Lys Gly Leu Leu Gly Met Pro Phe Asn Thr Thr Ala Ile Asp Gly Leu | | | | |
| 545 | 550 | 555 | 560 | |
| tat tgt gtt ggc gat agt tgc ttc cca gga caa ggt gtt ata gct gta | | | | 1728 |
| Tyr Cys Val Gly Asp Ser Cys Phe Pro Gly Gln Gly Val Ile Ala Val | | | | |
| 565 | 570 | 575 | | |
| gcc ttt tca gga gta atg tgc gct cat cgt gtt gca gct gac tta ggg | | | | 1776 |
| Ala Phe Ser Gly Val Met Cys Ala His Arg Val Ala Ala Asp Leu Gly | | | | |
| 580 | 585 | 590 | | |
| ttt gaa aaa aaa tca gat gtg ctg gac agt gct ctt ctt aga cta ctt | | | | 1824 |
| Phe Glu Lys Lys Ser Asp Val Leu Asp Ser Ala Leu Leu Arg Leu Leu | | | | |
| 595 | 600 | 605 | | |
| ggg tgg tta agg aca cta gca tga | | | | 1848 |
| Gly Trp Leu Arg Thr Leu Ala | | | | |
| 610 | 615 | | | |

<210> 76

<211> 615

<212> PRT

<213> Lycopersicon esculentum

<400> 76

| | | | |
|---|---|----|----|
| Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr | | | |
| 1 | 5 | 10 | 15 |

| | | | |
|---|----|----|--|
| Cys Lys Thr Val Ala Leu Gly Asp Ser Lys Pro Arg Tyr Asn Lys Gln | | | |
| 20 | 25 | 30 | |

| | | | |
|---|----|----|--|
| Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln | | | |
| 35 | 40 | 45 | |

Gln Gln Leu Cys Gly Leu Ser Trp Gly Val Asp Lys Ala Lys Gly Arg
50 55 60

Arg Gly Gly Thr Val Ser Asn Leu Lys Ala Val Val Asp Val Asp Lys
65 70 75 80

Arg Val Glu Ser Tyr Gly Ser Ser Asp Val Glu Gly Asn Glu Ser Gly
85 90 95

Ser Tyr Asp Ala Ile Val Ile Gly Ser Gly Ile Gly Gly Leu Val Ala
100 105 110

Ala Thr Gln Leu Ala Val Lys Gly Ala Lys Val Leu Val Leu Glu Lys
115 120 125

Tyr Val Ile Pro Gly Gly Ser Ser Gly Phe Tyr Glu Arg Asp Gly Tyr
130 135 140

Lys Phe Asp Val Gly Ser Ser Val Met Phe Gly Phe Ser Asp Lys Gly
145 150 155 160

Asn Leu Asn Leu Ile Thr Gln Ala Leu Ala Ala Val Gly Arg Lys Leu
165 170 175

Glu Val Ile Pro Asp Pro Thr Thr Val His Phe His Leu Pro Asn Asp
180 185 190

Leu Ser Val Arg Ile His Arg Glu Tyr Asp Asp Phe Ile Glu Glu Leu
195 200 205

Val Ser Lys Phe Pro His Glu Lys Glu Gly Ile Ile Lys Phe Tyr Ser
210 215 220

Glu Cys Trp Lys Ile Phe Asn Ser Leu Asn Ser Leu Glu Leu Lys Ser
225 230 235 240

Leu Glu Glu Pro Ile Tyr Leu Phe Gly Gln Phe Phe Lys Lys Pro Leu
245 250 255

Glu Cys Leu Thr Leu Ala Tyr Tyr Leu Pro Gln Asn Ala Gly Ser Ile
260 265 270

Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala
275 280 285

Glu Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gln Thr Pro Met Ile
290 295 300

Asn Ala Ser Met Val Leu Cys Asp Arg His Phe Gly Gly Ile Asn Tyr
305 310 315 320

Pro Val Gly Gly Val Gly Glu Ile Ala Lys Ser Leu Ala Lys Gly Leu
325 330 335

Asp Asp His Gly Ser Gln Ile Leu Tyr Arg Ala Asn Val Thr Ser Ile
340 345 350

Ile Leu Asp Asn Gly Lys Ala Val Gly Val Lys Leu Ser Asp Gly Arg
355 360 365

Lys Phe Tyr Ala Lys Thr Ile Val Ser Asn Ala Thr Arg Trp Asp Thr
370 375 380

Phe Gly Lys Leu Leu Lys Ala Glu Asn Leu Pro Lys Glu Glu Glu Asn
385 390 395 400

Phe Gln Lys Ala Tyr Val Lys Ala Pro Ser Phe Leu Ser Ile His Met
405 410 415

Gly Val Lys Ala Asp Val Leu Pro Pro Asp Thr Asp Cys His His Phe
420 425 430

Val Leu Glu Asp Asp Trp Thr Asn Leu Glu Lys Pro Tyr Gly Ser Ile
435 440 445

Phe Leu Ser Ile Pro Thr Val Leu Asp Ser Ser Leu Ala Pro Glu Gly
450 455 460

His His Ile Leu His Ile Phe Thr Thr Ser Ser Ile Glu Asp Trp Glu
465 470 475 480

Gly Leu Ser Pro Lys Asp Tyr Glu Ala Lys Lys Glu Val Val Ala Glu
485 490 495

147

Arg Ile Ile Ser Arg Leu Glu Lys Thr Leu Phe Pro Gly Leu Lys Ser
500 505 510

Ser Ile Leu Phe Lys Glu Val Gly Thr Pro Lys Thr His Arg Arg Tyr
515 520 525

Leu Ala Arg Asp Ser Gly Thr Tyr Gly Pro Met Pro Arg Gly Thr Pro
530 535 540

Lys Gly Leu Leu Gly Met Pro Phe Asn Thr Thr Ala Ile Asp Gly Leu
545 550 555 560

Tyr Cys Val Gly Asp Ser Cys Phe Pro Gly Gln Gly Val Ile Ala Val
565 570 575

Ala Phe Ser Gly Val Met Cys Ala His Arg Val Ala Ala Asp Leu Gly
580 585 590

Phe Glu Lys Lys Ser Asp Val Leu Asp Ser Ala Leu Leu Arg Leu Leu
595 600 605

Gly Trp Leu Arg Thr Leu Ala
610 615

<210> 77

<211> 1233

<212> DNA

<213> Tagetes erecta

<220>

<221> CDS

<222> (1) ..(1233)

<223>

<400> 77
atg gcc aca cac aaa ctc ctt caa ttc acc acc aat ctc cca cca tct 48
Met Ala Thr His Lys Leu Leu Gln Phe Thr Thr Asn Leu Pro Pro Ser
1 5 10 15

| | |
|---|-----|
| tct tct tca atc tct act ggc tgt tca ctc tcc ccc ttc ttc ctc aaa Ser Ser Ser Ile Ser Thr Gly Cys Ser Leu Ser Pro Phe Phe Leu Lys 20 25 30 | 96 |
| tca tct tct cat tcc cct aac cct cgc cga cac cgc cgc tcc gcc gta Ser Ser Ser His Ser Pro Asn Pro Arg Arg His Arg Arg Ser Ala Val 35 40 45 | 144 |
| tgc tgc tct ttc gcc tca ctc gac tct gca aaa atc aaa gtc gtt ggc Cys Cys Ser Phe Ala Ser Leu Asp Ser Ala Lys Ile Lys Val Val Gly 50 55 60 | 192 |
| gtc ggt ggt ggt ggc aac aat gcc gtt aac cgc atg att ggt agc ggc Val Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Gly Ser Gly 65 70 75 80 | 240 |
| tta cag ggt gtt gat ttt tac gcc att aac acg gac tca caa gcg ctt Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu 85 90 95 | 288 |
| ctg caa tct gtt gca cat aac cct att caa att ggg gag ctt ttg act Leu Gln Ser Val Ala His Asn Pro Ile Gln Ile Gly Glu Leu Leu Thr 100 105 110 | 336 |
| cgt gga tta ggt act ggt ggg aac ccg ctt ttg gga gaa cag gct gcg Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala 115 120 125 | 384 |
| gag gag tcg aag gaa gcg att ggg aat gcg ctt aaa ggg tcg gat ctt Glu Glu Ser Lys Glu Ala Ile Gly Asn Ala Leu Lys Gly Ser Asp Leu 130 135 140 | 432 |
| gtg ttt ata aca gca ggt atg ggt ggg acg ggt tcg ggt gct gct Val Phe Ile Thr Ala Gly Met Gly Gly Thr Gly Ser Gly Ala Ala 145 150 155 160 | 480 |
| cca gtt gta gcg cag ata gcg aaa gaa gca ggg tat tta act gtt ggt Pro Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly 165 170 175 | 528 |
| gtt gta acg tac cca ttc agc ttt gaa ggc cgt aaa aga tca gta cag Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln 180 185 190 | 576 |
| gcg tta gag gct att gag aag ctg caa aag aac gtt gac aca ctt ata Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile 195 200 205 | 624 |
| gtg att cca aat gac cgt ttg ctg gat att gct gat gaa aac acg cct Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Asn Thr Pro 210 215 220 | 672 |
| ctt cag gat gct ttt ctt ctt gct gat gat gta ctc cgc caa gga gtt Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val 225 230 235 240 | 720 |

| | |
|---|------|
| caa gga atc tca gat ata att aca ata cct ggg ctg gta aat gtg gac Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp 245 250 255 | 768 |
| ttt gca gac gtt aaa gca gtc atg aaa gat tct gga act gca atg ctt Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu 260 265 270 | 816 |
| ggt gtc ggt gtt tcc tca agt aaa aac cga gct gaa gaa gca gct gaa Gly Val Gly Val Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu 275 280 285 | 864 |
| caa gca act ctt gct cct ttg att gga tca tca att caa tct gct aca Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr 290 295 300 | 912 |
| ggt gtt gtt tat aat att acc gga ggg aag gac ata act cta caa gaa Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu 305 310 315 320 | 960 |
| gtc aac agg gtt tct cag gtg gta aca agt ttg gca gat cca tca gca Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala 325 330 335 | 1008 |
| aac att ata ttc ggg gca gtg gta gat gag aga tac aac ggg gag att Asn Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile 340 345 350 | 1056 |
| cat gtg acc att gtt gct act ggc ttt gcc cag tcg ttt cag aaa tct His Val Thr Ile Val Ala Thr Gly Phe Ala Gln Ser Phe Gln Lys Ser 355 360 365 | 1104 |
| ctt ctt gct gac ccg aaa gga gca aaa ctt gtt gat aga aat caa gaa Leu Leu Ala Asp Pro Lys Gly Ala Lys Leu Val Asp Arg Asn Gln Glu 370 375 380 | 1152 |
| cct aca caa cct ttg act tcc gcg aga tct ttg aca aca cct tct cct Pro Thr Gln Pro Leu Thr Ser Ala Arg Ser Leu Thr Thr Pro Ser Pro 385 390 395 400 | 1200 |
| gct ccg tct cgg tct agg aaa ctc ttc ttt taa Ala Pro Ser Arg Ser Arg Lys Leu Phe Phe 405 410 | 1233 |

<210> 78

<211> 410

<212> PRT

<213> Tagetes erecta

<400> 78

Met Ala Thr His Lys Leu Leu Gln Phe Thr Thr Asn Leu Pro Pro Ser
1 5 10 15

Ser Ser Ser Ile Ser Thr Gly Cys Ser Leu Ser Pro Phe Phe Leu Lys
20 25 30

Ser Ser Ser His Ser Pro Asn Pro Arg Arg His Arg Arg Ser Ala Val
35 40 45

Cys Cys Ser Phe Ala Ser Leu Asp Ser Ala Lys Ile Lys Val Val Gly
50 55 60

Val Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Gly Ser Gly
65 70 75 80

Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu
85 90 95

Leu Gln Ser Val Ala His Asn Pro Ile Gln Ile Gly Glu Leu Leu Thr
100 105 110

Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala
115 120 125

Glu Glu Ser Lys Glu Ala Ile Gly Asn Ala Leu Lys Gly Ser Asp Leu
130 135 140

Val Phe Ile Thr Ala Gly Met Gly Gly Thr Gly Ser Gly Ala Ala
145 150 155 160

Pro Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly
165 170 175

Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln
180 185 190

Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile
195 200 205

Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Asn Thr Pro
210 215 220

Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val
225 230 235 240

Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp
245 250 255

Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu
260 265 270

Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu
275 280 285

Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr
290 295 300

Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu
305 310 315 320

Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala
325 330 335

Asn Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile
340 345 350

His Val Thr Ile Val Ala Thr Gly Phe Ala Gln Ser Phe Gln Lys Ser
355 360 365

Leu Leu Ala Asp Pro Lys Gly Ala Lys Leu Val Asp Arg Asn Gln Glu
370 375 380

Pro Thr Gln Pro Leu Thr Ser Ala Arg Ser Leu Thr Thr Pro Ser Pro
385 390 395 400

Ala Pro Ser Arg Ser Arg Lys Leu Phe Phe
405 410

<210> 79

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer_bind

<222> (1)..(33)

<223>

<400> 79

gcatgctcta gaccttataa agatattttg tga

33

<210> 80

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer_bind

<222> (1)..(33)

<223>

<400> 80

gcatgcatct agaaatggtt cagtgtcaac cat

33

<210> 81

<211> 805

<212> DNA

<213> Nostoc sp. Strain PCC7120

<220>

<221> variation

<222> (1)..(805)

<223>

<400> 81
gcatgcacatc agaaatgggt cagtgtcaac catcatctct gcattcagaa aaactggtgt 60
tattgtcatac gacaatcaga gatgataaaa atattaataa gggtatattt attgcctgct 120
ttatcttatt tttatgggca attagtttaa tcatttact ctcaatagat acatccataa 180
ttcataagag cttatttaggt atagccatgc tttggcagac cttcttatat acaggtttat 240
ttattactgc tcatgtatgcc atgcacggcg tagtttatcc caaaaatccc agaataaata 300
attttataagg taagctcaact ctaatcttgt atggactact cccttataaa gatttattga 360
aaaaaacattg gttacaccac ggacatcctg gtactgattt agaccctgat tattacaatg 420
gtcatccccca aaacttcttt ctttggtatac tacattttat gaagtcttat tggcgatgga 480
cgcaaatttt cggatttagtg atgatttttc atggacttaa aaatctggtg catataccag 540
aaaataattt aattatattt tggatgatac cttcttattt aagttcagta caactatttt 600
attttggtagt atttttgctt cataaaaaagc tagaagggtgg ttatactaac ccccattgtg 660
cgcgcagtat cccattaccc cttttttggc cttttgttac ttgttatcac ttccggctacc 720
acaaggaaca tcacgaatac cctcaacttc cttggtgaa attacctgaa gctcacaaaa 780
tatctttata aggtcttagag catgc 805

<210> 82

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(37)

<223>

<400> 82
gcgcatgcat ctagaaatga tccagttaga acaaccca 37

<210> 83

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(37)

<223>

<400> 83
gcgcatgctc tagactatgg tgctttgtaa atttctg 37

<210> 84

<211> 792

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220> [REDACTED]

<221> misc_feature

<222> (5)..(775)

<223>

<400> 84
gcgcatgcat ctagaaatga tccagttaga acaaccactc agtcatcaag caaaaactgac 60
tccagtaactg agaagtaaat ctcagttaa ggggctttc attgctattg tcattgttag 120
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 180
gatgttattg cctgttatac tatggcaaac atttttatat acgggattat ttattacatc 240

tcatgatgcc atgcattggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg 300
aacattgacc ctatcccttt atggctttt accatatcaa aaactattga aaaaacattg 360
gttacaccac cacaatccag caagctcaat agacccggat tttcacaatg gtaaacacca 420
aagtttcttt gcttggattt ttcattttat gaaaggttac tggagttggg ggcaaataat 480
tgcgttgact attatttata actttgctaa atacatactc catatccaa gtgataatct 540
aacttacttt tgggtgctac cctcgctttt aagttcatta caattattct attttggcac 600
tttttaccc catagtgaac caataggggg ttatgttcag cctcattgtg cccaaacaat 660
tagccgtcct atttgggtt catttatacac gtgctatcat tttggctacc acgaggaaca 720
tcacgaatat cctcatattt cttggggca gttaccagaa atttacaaag caaaatagtc 780
tagagcatgc gc 792

<210> 85

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)...(26)

<223>

<400> 85

gtcgaccctg ctttaatgag atatgc

26

<210> 86

<211> 27

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(27)

<223>

<400> 86

ctcgagcttg gacaatcagt aaattga

27

<210> 87

<211> 210

<212> DNA

<213> Agrobacterium tumefaciens

<220>

<221> Terminator

<222> (1)..(210)

<223>

<400> 87

gtcgaccctg ctttaatgag atatgcgaga cgccatatgtat cgcatgatat ttgcttcaa 60

ttctgttgtg cacgttgtaa aaaacctgag catgtgttagc tcagatcctt accggccggtt 120

tcggttcatt ctaatgaata tatcacccgt tactatcgta ttttatgaa taatattctc 180

cgttcaattt actgattgtc caagctcgag 210

<210> 88

<211> 1773

<212> DNA

<213> Petunia hybrida

<220>

<221> Promotor
<222> (1)..(1773)
<223>

tttgaatagt cgactacatt ggataatgga acttctcgaa ttaacttcga attagtcgag 1500
ccccaaaataa tatatacgtc gggtgaaaaa ctataaaatg tttgacaaaaa atgtcaaatt 1560
aatatatcaa tctgcaacaa cctttcacc ttgagaacac agctgaaaatt ttttacaaag 1620
gtagttggtg aagctagtca gcgaatccca ttaccttcca ctctacctaa cccccttcac 1680
caacaacaaa tttctgtaat taaaaacta gccaaaaaaaaag aactctcttt tacaaagagc 1740
caaagactca atctttactt tcaagaaaaag ctt 1773

<210> 89

<211> 29

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(29)

<223>

<400> 89

gagctcttagc gcaatcttat gtggtacaa

29

<210> 90

<211> 29

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(29)

<223>

<400> 90
aagctttct tgaaagtaaa gattgagtc

29

<210> 91

<211> 39

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (39)

<223>

<400> 91
gcgcatgcat ctagaaatga attttgtga taaaccagt

39

<210> 92

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (37)

<223>

<400> 92
gcgcatgctc tagattacga attggttact gaattgt

37

<210> 93

<211> 819

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> misc_feature

<222> (5) .. (802)

<223>

<400> 93
gcgcatgcat ctagaaatga atttttgtga taaaccagtt agctattatg ttgcaataga 60
gcaatattaagt gctaaagaag atactgtttg ggggctggtg attgtcatag taattattag 120
tctttggta gctagttgg ctttttact agctattaat tatgccaaag tcccaatttg 180
gttgataacctt attgcaatag tttggcaaattt gttcctttat acagggctat ttattactgc 240
acatgatgct atgcatgggt cagtttatcg taaaaatccc aaaattaata attttatcgg 300
ttcacatagct gtagcgcttt acgctgtgtt tccatatcaa cagatgttaa agaattcattg 360
cttacatcat cgtcatcctg ctagcgaagt tgacccagat tttcatgatg gtaagagaac 420
aaacgctatt ttctggtatac tccatttcata gatagaatac tccagttggc aacagttaat 480
agtactaact atcctattta attagctaa atacgttttgc cacatccatc aaataaatct 540
catcttattt tggagtattc ctccaaatttt aagttccatt caactgtttt attcggAAC 600
atttttgcct catcgagaac ccaagaaagg atatgtttat ccccatggca gccaaacaat 660
aaaattgcca actttttgtt catttatcgat ttgctaccac tttggttatac atgaagaaca 720
tcatgagtat ccccatgtac cttggggca acttccatct gtatataagc agagagtatt 780
caacaattca gtaaccaatt cgtaatctag agcatgcgc 819

<210> 94

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (33)

<223>

<400> 94

gcgcatgcat ctagaaatgg cgatcgccat tat

33

<210> 95

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (32)

<223>

<400> 95

gcgcatgctc tagatcacaa atttgattta ga

32

<210> 96

<211> 720

<212> DNA

<213> Nodularia spumigena NSOR10

<220>

<221> CDS

<222> (5) .. (703)

<223>

<400> 96
 gcgc atg cat cta gaa atg gcg atc gcc att att agt ata tgg gct atc 49
 Met His Leu Glu Met Ala Ile Ala Ile Ser Ile Trp Ala Ile
 1 5 10 15

agc cta ggt ttg tta ctt tat att gat ata tcc caa ttc aag ttt tgg 97
 Ser Leu Gly Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp
 20 25 30

atg ttg tta ccg ctc ata ttt tgg caa aca ttt tta tat acg gga tta 145
 Met Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu
 35 40 45

ttt att aca gct cat gat gcc atg cat ggg gta gtt ttt ccc aaa aat 193
 Phe Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn
 50 55 60

ccc aaa atc aac cat ttc att ggc tca ttg tgc ctg ttt ctt tat ggt 241
 Pro Lys Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly
 65 70 75

ctt tta cct tat caa aaa ctt tta aaa aag cat tgg cta cat cac cat 289
 Leu Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His His
 80 85 90 95

aat cca gcc agt gaa aca gat cca gat ttt cac aac ggg aag cag aaa 337
 Asn Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys
 100 105 110

aac ttt ttt gct tgg tat tta tat ttt atg aag cgt tac tgg agt tgg 385
 Asn Phe Phe Ala Trp Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp
 115 120 125

tta caa att atc aca tta atg att att tat aac tta cta aaa tat ata 433
 Leu Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile
 130 135 140

tgg cat ttt cca gag gat aat atg act tat ttt tgg gta gtt ccc tca 481
 Trp His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser
 145 150 155

att tta agt tct tta caa tta ttt tat ttt gga act ttt cta ccc cac 529
 Ile Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His
 160 165 170 175

agt gag cct gta gaa ggt tat aaa gag cct cat cgt tcc caa act att 577
 Ser Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile
 180 185 190

agc cgt ccc att tgg tgg tca ttt ata act tgt tac cat ttt ggt tat 625
 Ser Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr
 195 200 205

cat tac gaa cat cat gaa tac ccc cat gtt cct tgg tgg caa tta cca His Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro
210 215 220

gaa att tat aaa atg tct aaa tca aat ttg tgatctagag catgcgc 720
 Glu Ile Tyr Lys Met Ser Lys Ser Asn Leu
 225 230

<210> 97

<211> 233

<212> PRT

<213> Nodularia spumigena NSOR10

<400> 97

| | | | |
|-----------------|--------------------------------------|-------------|----|
| Met His Leu Glu | Met Ala Ile Ala Ile Ile Ser Ile Trp. | Ala Ile Ser | |
| 1 | 5 | 10 | 15 |

Leu Gly Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met
..... 20 25 30

Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe
 35 40 45

Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro
50 55 60

Lys Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu
65 70 75 80

Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His His Asn
85 90 95

Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn
100 105 110

Phe Phe Ala Trp Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu
115 120 125

Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp
130 135 140

His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile
145 150 155 160

Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser
165 170 175

Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser
180 185 190

Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His
195 200 205

Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu
210 215 220

Ile Tyr Lys Met Ser Lys Ser Asn Leu
225 230

<210> 98

<211> 37

<212> DNA

<213> Künstliche Sequenz

{}

<220>

<221> primer

<222> (1)...(37)

<223>

<400> 98

gagctcgata tctttgccag tattacaaca gcttata

37

<210> 99

<211> 31

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(31)

<223>

<400> 99

cccggttta ctgaaaaata acagtaaaac c

31

<210> 100

<211> 2096

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(2096)

<223>

<400> 100

gagctcgata tctttgccag tattacaaca gcttatatgt tgaggcaggta aaagcttcaa

60

tgccctattc tttctacagt tatcaatgtt gctcgctaa tatctggtgt tcttcgcga

120

atgtcaattg gcttgcagca cattgtcctc taatatccat tcaagcttct tagatgatga

180

aacatttgtc aaatttatta atttcatagt gttcagtctc aattctttag ctggttcctc

240

atagtaaagt tgtctaataat gaaatgaaaa tttctgtgt gttgtactaa tacctttca

300

tggttgtcta tagaacgtcg atgaagagcc aaacagaaac tattttggc tgcgatttct

360

gataccattg tatctgaatg ctgggtggga gctcatcaga agctttacaa tgggtcacat

420

atatggagcc ggtatgagga atgctggaa tcagttgcgt ttcgcgtgct aggactttc

480

cttcctggta tttctgccc aagccagtt gattacgtga actccgtcag acttggaaag

540

<211> 24

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(24)

<223>

<400> 101

gagctcgaat tctctgaaaa ggag

24

<210> 102

<211> 25

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(25)

<223>

[]

<400> 102

aagctttata gagaatgtat aagat

25

<210> 103

<211> 1222

<212> DNA

[] []

<213> Lycopersicon esculentum

[]

<220>

<221> Promotor

<222> (1)..(1222)

<223>

<400> 103

gagctcgaat tctctgaaaa ggaggaccat atttgcgcga ctgtggttca tatttccaag 60
tacattttaga tgaactatacatcagattt aaggttattt gtataatcaa tccagtggat 120
tctcgttctg gcaccccttag aagtacatgt gcggaaaaga atgataaggt ttgtattt 180
gttgacaaag cctgttgcct ttctcattt taaatgttct gaacgactcc taaattactc 240
ttaaggtgtt aggtcttccg tgcctgtttt taaatataat gctgtgccgt gacttacctt 300
ttgttaccatt tgttcaaatg tatggcctga acaccagggt tgtcaaaaat gtctcatgcc 360
cgttttattt gtcgtaaaaat ggcgtgatgc caaattctgc cgctccacag tgagcatttc 420
gatctactgg aaattgacca acttattttt tcacttgata actaaacaaa atcctattaa 480
ctttaatcat acattgttattt tataccgaaa aatttatgca taactcatta aattaccttt 540
tttagcagtc aaattctaaa tcagtttcta atttatcaaa atggctttt tagggtccca 600
tttccactaa tatacctgcc gtccatgcac tgactacaaa acaaataacct cactatgttt 660
gttagtgctt ggtaatataa aacctttctt tttatgagaa agttcaccga gaataatttt 720
ctatttgtgg cataatagta tatagtgcag attgacaaga atttaatttt gcagttggc 780
acatgaacaa tttcctcaa agttgttagaa agtacttttcaattttctt caccgaaaat 840
tatttataat tgaaattaaa accgaatgag ctgcaagattt caagtcgaat tttcaaaaaga 900
attgaccaag aaaaaattca aaaatatccc ccacccctta ccaaacacat cctaaagtga 960
ggtagatgact gggactggga ttggaaaaag ggtaaaaatgc tttcactagc ttagcaaaga 1020
ttccactttt ttagctatct ttctttctca tttccttttt tctttttctt ttttttgtt 1080
tataagccaa agtaggtacc caaaagcatc aatattttgtt attgcttgggt gattcctctg 1140
tagtccagta tttcattttc tacaagttcc acctccctcc ataattaacc attatcaatc 1200
ttatacatc tctataaagc tt 1222

<210> 104

<211> 30

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (30)

<223>

<400> 104

gagctctaca aattagggtt actttattca

30

<210> 105

<211> 32

: : /

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (32)

<223>

<400> 105

aagcttagaa ttatccaa attccgtaaa tc

32

<210> 106

<211> 1540

: : /

<212> Cucumis sativus

<220>

<221> Promotor

<222> (1)..(1540)

<223>

<400> 106
gagctctaca aattagggtt actttattca ttttcatcca ttctctttat tgtaaaat 60
tgtacattta ttcaataata ttatatgttt attacaaatt ctcactttct tattcatacc 120
tattcactca agcctttacc atcttccttt tctatcaa tactatttct acttcatttt 180
tcacgaaaaaacatctttc tttatattctt gtccacttcg ttttagggatg cctaattgtcc 240
caaatttcat ctctcgtagt aacacaaaac caatgtaatg ctacttctct ctacatttt 300
aatacaaaata aagtgaaaca aaatatctat aaataaaacaa atatataat tttgttagac 360
gctgtctcaa cccatcaatt aaaaaatttt gttatatttc tacatttacct actaaatttg 420
tttctcatat ttacctttt accccccacaa aaaaaaatta taaaaaaagaa agaaaaaaagc 480
taaaccctat ttaaatagct aactataaga tcttaaaattt atcctcatca gtgtatagtt 540
taattggta ttaacttata acattatata tctatgacat atactctctc ctagctattt 600
ctcacatttt ttaacttaag aaaatagtca taacatagtc taaaattcaa acatccacat 660
gctctaattt gattaacaaa aagttagaaa tattttat taaaattttt aataaaaaag actaataat 720
atataaaatg aatgttcata cgcaaaaaa ttttagagatg agtatgcttt cacatgctga 780
gattatttc aaaactaagg ttgttagcaat attaaatcaa taaaattttt ataaataaca 840
aaattaacct gctcggttt gctgtatag ggaggctaca aaataaaatta aactaaagat 900
gattatgttt tagacatttt ttctatctgt attagtttat acatattaat tcaggagctg 960
cacaacccaa ttctatccc gttcccttgggt ggctgggtt ctcacaagg tcaatagtca 1020
atattaggtt ttattggact ttaatagta tcaaacaaaat ctatgtgtga actaaaaat 1080
tgtattaaat atttagggta acctgttgcc gtttttagaa taatgtttct tcttaataca 1140
cgaaagcgta ttgtgtattc attcatttgg cgccctcacat gcttcgggtt gctcgcttta 1200
gtctctgcct tctttgtata ttgtactccc cctcttccta tgccacgtgt tctgagctta 1260
acaagccacg ttgcgtgccaa tggccaaaca agtcattttt acttcacaag gtccgatttg 1320
acctccaaaaa caacgacaag tttccgaaca gtcgcgaaga tcaagggtat aatcgcttt 1380

ttgaattcta tttctcttta tttaatagtc cctctcggt gatagtttt aaaagattt 1440
taaaaacgtag ctgctgttta agtaaatccc agtccttcag ttttgcttt tgtgtgttt 1500
gtttctctga ttacggaat ttggaaataa ttctaagctt 1540